

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 03-24-04
Searcher: Reverly C. 2528
Terminal time: 201
Elapsed time: _____
CPU time: _____
Total time: 28
Number of Searches: _____
Number of Databases: 1

Search Site
____ STIC
____ CM-1
____ Pre-S
Type of Search
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors
____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other CGN

```
QY 1741 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTCAATCTGTAATCGAGATCATCCAGTC 1800
Db 1741 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTCAATCTGTAATCGAGATCATCCAGTC 1800
QY 1801 GAATCCTCAAGGAATAATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCC 1860
Db 1801 GAATCCTCAAGGAATAATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCC 1860
QY 1861 TCCACAGTTGCTATCCTTATGTCGAAGTCTTAATGACTCTACACAGAGTTTCATAGC 1920
Db 1861 TCCACAGTTGCTATCCTTATGTCGAAGTCTTAATGACTCTACACAGAGTTTCATAGC 1920
QY 1921 TAAATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTC 1980
Db 1921 TAAATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTC 1980
QY 1981 CTTTGAAGAACTCCTGCTGTTTATTCAGCAGTGCTGCTCAATATCGGCTCTATATTC 2040
Db 1981 CTTTGAAGAACTCCTGCTGTTTATTCAGCAGTGCTGCTCAATATCGGCTCTATATTC 2040
QY 2041 TGGTATCTCCAAATACACACACTAGCTTTCGTTGGAGAGATTATTATGCTTGGT 2100
Db 2041 TGGTATCTCCAAATACACACACTAGCTTTCGTTGGAGAGATTATTATGCTTGGT 2100
QY 2101 AAGGCTTTGTTGAGGCTTACCAACACACTAGGAGATCTTCAATAAATAAAGA 2156
Db 2101 AAGGCTTTGTTGAGGCTTACCAACACACTAGGAGATCTTCAATAAATAAAGA 2156

RESULT 2
AE001654/c
LOCUS      10421 bp      DNA      linear      BCT 01-DEC-2000
          amydia pneumoniae section 70 of 103 of the complete genome.
          91654 AE001363
          91654.1 GI:4377031

-----amydophila pneumoniae CWL029
Bacteria: Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
1 (bases 1 to 10421)
Kalan, S., Mitchell, W., Marathe, R., Lammel, C., Pan, J., Hyman, R.W.,
Olinger, L., Grimwood, J., Davis, R.W. and Stephens, R.S.
Comparative genomes of Chlamydia pneumoniae and C. trachomatis
Nat. Genet. 21 (4), 385-389 (1999)
99206606
PUBMED 10192388
2 (bases 1 to 10421)
Kalan, S., Mitchell, W., Marathe, R., Lammel, C., Pan, J., Olinger, L.,
Grimwood, J., Davis, R.W. and Stephens, R.S.
Direct Submission
Submitted (01-DEC-1998) Program in Infectious Diseases, University
of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA
Location/Qualifiers
  1. .10421
    /organism="Chlamydia pneumoniae CWL029"
    /mol_type="genomic DNA"
    /strain="CWL029"
    /db_xref="taxon:115713"
    133..2751
    /genes="Cp0727"
    133..2751
    /genes="Cp0727"
    /codon_start=1
    /transl_table=11
    /product="Cp619 hypothetical protein"
    /protein_id="AA018866.1"
    /db_xref="GI:4377032"

gene
CDS
  133..2751
    /genes="Cp0727"
    /codon_start=1
    /transl_table=11
    /product="Integral Membrane Protein"
    /protein_id="AA018869.1"
    /db_xref="GI:4377035"
    /translation="MSRDKNEVLSARIFNILSGTCSRITGITPREIAMATYFGADPI
    VAAFWLGRFTVFRLKILGLILEQAFIPHFELRAQSLDRAAFRRFRRLIKGSTI
    IFILLIEAVLWVLYQVEGTDMILLTLLPCGIFLMNVNNGALHCHENKFFGVG
    LAPVNNIWIFFVIAARHSDPRRIIGLSVALVYGFEEWLTIVPGVWKFLLEAKSP
    POEHSVRALLAPLSGLTSSIFOLNLSIDICLARYHEITGLVLYMSLKIYOLPIH
    LFPGCVTVLLPAISRCVQREHERGLKMKFVLTLVMSVIMTAGLILLALPGVRV
    LYEHGLFPQSAVAVLVRLGYGAIIPMALPLVSVLVFVQYAVPLFIFIGITALA
    NIVLSLGRWLKDSIGSYATISITAWQVLYFWYSSKRLPMYSLKLMESIRRSIK
    VMGTMLACMTLGLNTLTQTVYFNLNLTPLAWPLSSITAAQIAFLSBSCLPLAFL
    FGAKLLRVBDLINLASFEYWRGQGLQCHVMDTCN"
    8143..8319
    /gene="Cp0731"
    8143..8319
    /gene="Cp0731"

LOADDIAIKAPALGGQDEIITLQTLVPLVDKTTFTKABFDLIYVTTATPNTASLKL
YLTRQIAEYRGKITKVYONSIONSETKVENNRSMLETLQSLMFOAOQCNFVYPIH
QANALNIAITFKVTSALTSMEMVYGLLGLSVYMERLADDEKXAIFFKSNVEYPIH
VGGSWNGHIAKVAAYQELAYSLSLGTAVTSQDQIKAYLQTRGNEFKATRRFFNIGD
QMOFATPFGNCTTANGAIOPDLGGPIREAMTVGTVEADIVSNQORLILNFNTA
ATAHVLQIQIAELQKADDDLPKASFTENRFKFAVAMITSSLDALISMLNSQ
LPGQEAFLKELIEINFNLAANALSLQITNEFTSVYSLSVISVSKTQNLFF
AGDYETLLAAREYVYRTARCKOALNVLGLOKINSLPGATSAQKQEMNALT
YXOYSLVTLNLTLESLAGLKWTLQTTSNKYDKSVFKIESFDWDWIPTLALESF
LTSFPFNISATGGLPIPTQVSDQIYTSQGTQQLNLQNMOTTIQEWTLVSTSMQ
VLNGILSQLAGAIYSN"
complement (3078..5033)
/gene="Cp0728"
complement (3078..5033)
/gene="Cp0728"
/codon_start=1
/transl_table=11
/product="CHLPN 76 kDa Homolog_1 (CT622)"
/protein_id="AA018867.1"
/db_xref="GI:4377033"
/translation="MVNPIGPGPIDETERTPPADLSAQGLEASANKSAEPAQIAGAE
AKPEKSDSVSRWSIILRSVNALMSLADKLGIASSNSSSTSTSAVDSTTATPTP
PPPTFDYKTAQATYDTITFTSLADIQAALVSLQDAVINIKDTAATDETAATAEW
ETKNADAVKGAQITELAKYASDNQALIDSLGKLTSLDLQALLOALQVANNKAAELL
KEMQDNVPVPGKTPAIQAQSLVDQDATQITQEDGNATRDAYFAGQNASGAVENAKSN
NSISNIDSAKAAIATATQIAEAQKFPDPSILOAEQMVIOAEKDLKNIKPADGSDV
PNPCTTVGSGKQGGSSIGSIRVSMLLDDENETASIIIMSGPROHMENTENPSQAA
QOELAAARAKAAGDSDSAAALADAQALBAALGKAGQOQGLNALQIASAAVSA
GVPPAAASSIGSSVKLYKTSKSTGSDYKQIISAGYDAIKSINDAYGRANDATRDVI
NNVSTPALTRSVPEARTAEKPEKTDQALARVISGNSRTLGDDVTSVALQSVNQIIQ
SNQANNEEIRQKLTSAVTPPOFGYPYVQLSNDSTQCTAKLESFLAEGSRTAAETK
ALSPTNLSLPIQVVLNIGLSYGLQ"
complement (5145..6413)
/gene="Cp0729"
complement (5145..6413)
/gene="Cp0729"
/codon_start=1
/transl_table=11
/product="CHLPN 76 kDa Homolog_2 (CT623)"
/protein_id="AA018868.1"
/db_xref="GI:4377034"
/translation="MKKQVQWLASVVLVLLATISYAEPLSEQKSHKSTVTTTDEVK
DYLKGEVETRKQDQVLRIRAGDVRAARWLFREDIKNPSDKDKYNPLFVNRYSEFYL
YIDYRABRNWLSKXNMTAAGGNTAGVDINRAFLGYRFPKNTPTDFPMFIEGRS
GLGLDFSEVQFQSNFGLHIYWTRELSDYQVYVHGGFVPMVMTKHHVAVVEGI
LNRLPKQFFVKCVSVNGLHVPFVSETSTTEKAAATKMKYKCVQWOLYKGSQVPAING
QKXELIYGAFLNPLAKATKTLNGHENLAWFPGGLGLRKGKADWSATVRYFVEA
LVSPEIDVSGIGRGNLLKFWFAIAAANYDPEANGFTNYKGFSAIYNGITDSLSFR
AYGAYKSPANDKLGSDFTRKFDIGIISAF"
complement (6516..8159)
/gene="mvin"
/notes="synonym: Cp0730"
complement (6516..8159)
/gene="mvin"
/codon_start=1
/transl_table=11
/product="Integral Membrane Protein"
/protein_id="AA018869.1"
/db_xref="GI:4377035"
/translation="MSRDKNEVLSARIFNILSGTCSRITGITPREIAMATYFGADPI
VAAFWLGRFTVFRLKILGLILEQAFIPHFELRAQSLDRAAFRRFRRLIKGSTI
IFILLIEAVLWVLYQVEGTDMILLTLLPCGIFLMNVNNGALHCHENKFFGVG
LAPVNNIWIFFVIAARHSDPRRIIGLSVALVYGFEEWLTIVPGVWKFLLEAKSP
POEHSVRALLAPLSGLTSSIFOLNLSIDICLARYHEITGLVLYMSLKIYOLPIH
LFPGCVTVLLPAISRCVQREHERGLKMKFVLTLVMSVIMTAGLILLALPGVRV
LYEHGLFPQSAVAVLVRLGYGAIIPMALPLVSVLVFVQYAVPLFIFIGITALA
NIVLSLGRWLKDSIGSYATISITAWQVLYFWYSSKRLPMYSLKLMESIRRSIK
VMGTMLACMTLGLNTLTQTVYFNLNLTPLAWPLSSITAAQIAFLSBSCLPLAFL
FGAKLLRVBDLINLASFEYWRGQGLQCHVMDTCN"
8143..8319
/gene="Cp0731"
8143..8319
/gene="Cp0731"
```


1561 TGCTTAAATCCATCAATGATGCTAGTGTAGGACGAAATGATGCGACTCGTGATG 1620
Db TGCTTAAATCCATCAATGATGCTAGTGTAGGACGAAATGATGCGACTCGTGATG 3514
Qy GATTAACATGTAAGTACCCCGCTCTACACGATCCGTTCTAGACGACGACGAGC 1680
Db GATTAACATGTAAGTACCCCGCTCTACACGATCCGTTCTAGACGACGACGAGC 3454
Qy TCGAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1740
Db TCGAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3394
Qy TCTTGGAGATGCTATAGTCAAGTTTCGCGACTCAATCTGTAATGACGATCAATC 1800
Db TCTTGGAGATGCTATAGTCAAGTTTCGCGACTCAATCTGTAATGACGATCAATC 3334
Qy GAATCTCAAGGATATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1860
Db GAATCTCAAGGATATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 3274
Qy TCCACAGTTTGGCTATCTTATGTCGAATCTTCAATGACTCTACACGAAAGTTCA 1920
Db TCCACAGTTTGGCTATCTTATGTCGAATCTTCAATGACTCTACACGAAAGTTCA 3214
Qy TAAATGAAAGTTTGGCTATCTTATGTCGAATCTTCAATGACTCTACACGAAAG 1980
Db TAAATGAAAGTTTGGCTATCTTATGTCGAATCTTCAATGACTCTACACGAAAG 3154
Qy CTTTGAACGAACTCTTGTATGTCGAATCTTCAATGACTCTACACGAAAGTTCA 2040
Db CTTTGAACGAACTCTTGTATGTCGAATCTTCAATGACTCTACACGAAAGTTCA 3094
Qy TGGTTATCTCCAATCAACACCTTGTATGTCGAATCTTCAATGACTCTACACGAA 2100
Db TGGTTATCTCCAATCAACACCTTGTATGTCGAATCTTCAATGACTCTACACGAA 3034
Qy AAGGCTTTGTTAGGCTTACCAACACCTTGTATGTCGAATCTTCAATGACTCTAC 2156
Db AAGGCTTTGTTAGGCTTACCAACACCTTGTATGTCGAATCTTCAATGACTCTAC 2978

RESULT 3
LOCUS AE002165 12548 bp DNA linear BCT 30-MAY-2000
DEFINITION Chlamydia pneumoniae AR39, section 2 of 94 of the complete genome.
ACCESSION AE002165 AE002161
VERSION AE002165.1 GI:7188948
KEYWORDS Chlamydia pneumoniae AR39
SOURCE Chlamydia pneumoniae AR39
ORGANISM Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.
REFERENCE 1 (bases 1 to 12548)
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Unayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, A., Craven, B., Bowman, C., Dodson, R., Gwin, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.
TITLE Genome sequences of Chlamydia trachomatis and Chlamydia pneumoniae AR39
JOURNAL Nucleic Acids Res. 28 (6), 1397-1406 (2000)
MEDLINE 20150255
FURNED 10684935
REFERENCE 2 (bases 1 to 12548)
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Unayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, A., Craven, B., Bowman, C., Dodson, R., Gwin, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

FEATURES
source

1. 12548
/organism="Chlamydia pneumoniae AR39"
/mol_type="genomic DNA"
/strain="AR39"
/db_xref="taxon:115711"
/notes="synonym: Chlamydia pneumoniae AR39"
297. .965
/gene="CP0011"
297. .965
/gene="CP0011"
/note="similar to SP:P31218 GB:X71492 PID:296947 GB:U00096
PID:1736770; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="uridine kinase"
/protein_id="AA037907.1"
/db_xref="GI:7188951"
/translation="MLMMLMIIGITGGGAGKTLTQNIKEIFGEDSVVICQDNYK
DRSHYTFEERANLIDHPDAFNDLLISDIKRLKNEIVQAPVDFVIGRSKTEIET
IYFSKVLVEGILVFENQELDLDIRIFVDIADERILERMVRDVOEGSDVDCIMS
RYLSMWKPEKEIEPTKRYADIIVHGNVRQNVVNIISQIKHLENALSDBTYYM
VNSK"
complement(1047. .2021)
/gene="CP0012"
complement(1047. .2021)
/gene="CP0012"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA037908.1"
/db_xref="GI:7188951"
/translation="MEKYYVALAYVITVDNPEHEIALHKKFLEDLDVCRIVISQ
GNGQSGYEPHAEIKMQLKRPNSKIKFHIHKNIFPRITVYKRELALGCE
VLSKQAKHLSPOEAKMEKLEKLECLLDVNNYEWKIGEDNATLPDIOTFRFPPEVA
EKLAQECDDPTTPMWTCTGICELYSVLEKKEVYQVGGVIAVQOQVGTGKX
LGKLFDDRLAIPIDSDPEVPIAECHCQCTPDSAYNCANTCNALFLCQDECIH
CHQCCGCEGSPRVRAFDSSRGKPFPRZHLCEISENSEASACCL1"
2238. .2867
/gene="CP0013"
2238. .2867
/gene="CP0013"
/note="similar to SP:P81288; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein S4"
/protein_id="AA037909.1"
/db_xref="GI:7188952"
/translation="MAYICGPKRVARRFGANIFGRSRNPLKKPPHPPGQGMQRKX
SGYGLQEEKQKLCACGMTEKQVKAPEVHKQCNVAQMELEFECRLDNNVYRM
DFATITFAAOLVAHGHLVNGRRVDRSPFLRPGMQLSLKSKSLQSLQVDALESKD
ESSLSYISLDTGTEKGLLAVSPQDQIEAQLPLPINISVVCFLSHRT"
complement(2986. .3867)
/gene="CP0014"
complement(2986. .3867)
/gene="CP0014"
/note="similar to SP:P12638 GB:M22591 PID:146954
PID:405898 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="endonuclease IV"
/protein_id="AA037910.1"
/db_xref="GI:7188953"
/translation="MKVLPSPPLICAHVSTAGGLKNAIYSGRDIGASTVQITFANQ
RQWRALKEVEDFKALKTDLISYIMSHAGLYLNGAPDVLLEKSRIGIYQIL
DCITLGSFYVHFHFGALKSKEDKNNKVSFQSPAPLSPFVLLVLETTAGGT
LIGNPFELGVQNLKNQIPIGCVDTCHIFAAGYDITSPQGWEDVINEFDEYVGLS
YLRAFHLNDSMFPLGANKRHAFLGEGYICKSFKLTMDTERTRKPKYLETFGGPN
WQKEIGELLAFSKNRDS"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:53:56 / Search time 17.7616 Seconds
(without alignments)
3525.629 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MNPPIGPGIDETERTPAD.....SLFIQQLVNLIGSLGYLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PTR 78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	2 D72042	conserved hypothet
2	3187	100.0	651	2 E86581	CHLPN 76 kDa homol
3	2205	89.2	715	2 I40729	hypothetical 76K p
4	899.5	28.2	647	2 G71490	hypothetical prote
5	232.5	7.3	2055	2 T31110	extracellular matr
6	232	7.3	1822	2 S33441	EF protein - strep
7	223.5	7.0	2481	2 D90011	FmtB protein [impo
8	221.5	7.0	971	2 B30835	probable tail fibe
9	221.5	7.0	973	2 C85693	probable membrane
10	220	6.9	6713	2 B99921	hypothetical prote
11	210	6.6	1122	2 G44887	probable tail fibe
12	210	6.6	2271	2 F90073	hypothetical prote
13	208.5	6.5	4776	2 E95206	cell wall surface
14	199	6.2	1365	2 T30822	Imp1 protein - Myc
15	196.5	6.2	1829	2 T24583	hypothetical prote
16	196.5	6.2	2232	2 T34434	hypothetical prote
17	192	6.0	1156	2 T34852	probable secreted
18	190.5	6.0	1063	2 D86731	hypothetical prote
19	189	5.9	2155	2 AD2742	conserved hypothet
20	189	5.9	2155	2 C97523	hypothetical prote
21	187	5.9	1147	2 T35781	hypothetical prote
22	186.5	5.9	1306	2 S25370	MSB2 protein - Yea
23	183	5.7	2186	2 H89960	hypothetical prote
24	182.5	5.7	641	2 C82206	methy1-accepting c
25	182.5	5.7	1238	2 T03465	probable exonuclea
26	182.5	5.7	2541	2 S11661	tailin - mouse
27	182	5.7	1269	2 F47330	probable myosin he
28	178.5	5.6	1566	2 A43607	cell surface antig
29	177	5.6	2090	2 S26058	probable transform

ALIGNMENTS

RESULT 1

D72042

conserved hypothetical protein CP0018 [imported] - Chlamydia pneumoniae (strains CW N:Alternate names: chlpn 76 kDa homolog 1 (cr622); hypothetical protein CPN0728

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C/Accession: D72042; D81623

R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A/Accession: D72042

A/Molecule type: DNA

A:Residues: 1-651 <REA>

A/Cross-references: GB:AE001654; GB:AE001363; NID:94377031; PIDN:AA018867.1; PID:943770:

A/Experimental source: strain CWL029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gyll, S.R.; Heideberg, J.F.; White, O.; Hick

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salz

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR3

A:Reference number: A81500; MUID:20150255; PMID:10684955

A/Accession: D81623

A>Status: preliminary

A/Molecule type: DNA

A:Residues: 1-651 <REA>

A/Cross-references: GB:AE002165; GB:AE002164; NID:57188948; PIDN:AAF37914.1; PID:571889:

A/Experimental source: strain AR39, HL cells

C/Comment: This sequence was originally identified as homologous to part of a sequence

PIR:H71490).

C/Genetics:

A:Gene: CPN0728; CP0018

Query Match

Best Local Similarity 100.0%; Score 3187; DB 2; Length 651;

Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNPPIGPGIDETERTPADLSAAGLEASAAKSAQAQIAGAAKPKESKTSVERWSI 60

Db 1 MNPPIGPGIDETERTPADLSAAGLEASAAKSAQAQIAGAAKPKESKTSVERWSI 60

Qy 61 LRSVAVNALSADKLGITASSNSSSTSRSDVDSTTATATPTPPPTFDYKTAQATYDT 120

Db 61 LRSVAVNALSADKLGITASSNSSSTSRSDVDSTTATATPTPPPTFDYKTAQATYDT 120

Qy 121 IFSTSLADIQAALVSLQDAVTNIKDTAETAETAAEWETKNADAVKGAQITELAKY 180

Db 121 IFSTSLADIQAALVSLQDAVTNIKDTAETAETAAEWETKNADAVKGAQITELAKY 180

Qy 181 ASDNQAILDSGLKTSFDLLQAALLOSVAANNKAELKEMQDNFVVGKTPAIAQSLVD 240

Db 181 ASDNQAILDSGLKTSFDLLQAALLOSVAANNKAELKEMQDNFVVGKTPAIAQSLVD 240

Db	301	KKFPDPSILOEAEQWVIOAEKLNKIKPAGDSVFNPGTTVGSGKQOQSGSIGSRVSMLL	360
Qy	361	DDAENETASTLMSGRQMIHMFTENTNPDSSQAQOEELAQARAAXAAGDSSAAALADAOX	420
Db	361	DDAENETASTLMSGRQMIHMFTENTNPDSSQAQOEELAQARAAXAAGDSSAAALADAOX	420
Qy	421	ALEAALKAGQOQOILNALGOIASAAVYAGVPPAAASSIGSSVKOLYKTSKSTGSDYKT	480
Db	421	ALEAALKAGQOQOILNALGOIASAAVYAGVPPAAASSIGSSVKOLYKTSKSTGSDYKT	480
Qy	481	QISAGYDAYKTSINDAYGEARNDATRDVINNVSTPALTRSVPRARTEARPEKTDQALRV	540
Db	481	QISAGYDAYKTSINDAYGEARNDATRDVINNVSTPALTRSVPRARTEARPEKTDQALRV	540
Qy	541	ISGSRITLGDVYSQVSALQSWQIISQNPQANNEERQKLTSAVTKPPQFGYPYVQLNSD	600
Db	541	ISGSRITLGDVYSQVSALQSWQIISQNPQANNEERQKLTSAVTKPPQFGYPYVQLNSD	600
Qy	601	STQKFIAXLESIFAEQSGSTAAEIKALSFETNSLFIQOVLWNIGSLYSGLYO	651
Db	601	STQKFIAXLESIFAEQSGSTAAEIKALSFETNSLFIQOVLWNIGSLYSGLYO	651

RESULT 3
I40729

601 STQKFIKLSLEAEGSGTAAEIKALSFETNSLFTQQVLNIGLSYGLQ 651
 |||||
 601 STQKFIKLSLEAEGSGTAAEIKALSFETNSLFTQQVLNIGLSYGLQ 651
 |||||
 DBB
 RESULT 2
 E86581
 CHLPN 76 kDa homolog 1 (CT622) [imported] - Chlamydomphila pneumoniae (strain J138)
 C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: E86581
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; IS
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
 A:Reference number: A86491. MUID:20930349; PMID:10871362

[illegible][illegible]

22

QY
121
TF1313HADIYATM1028

db / 377 IFTSTSLADIQAALVSLQDAVININDIHAIPDEEINRAAE

181 ASDNQAILDSLGKLTSEFDLLQALLQSVANNKAAELLNEMQDNEVFGRITLHQS

437 ASDNOAILDSLGLTSFDLLQALLQSVANNKAAELLKEMQDNFVVPFGKTPATAQSLVD 43

[illegible][illegible]

..... 360

100

Db
KKFPUSTTVEAEQNVIXATWZGAGS

QY 361 DDAENETASILMSGFRQMLHMFN TENFDSQHAQZELAAQACHTAQSDD

617 DDAENETASILMSGFRQMIHMFNTENPDSQAAQQLAAQARAAGAAGDDUSAAAHLLHDAQN 07

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	2204	100.0	651	2	D72042	conserved hypot
2	2204	100.0	651	2	E86581	CHLPP 76 kDa homol
3	2201	99.9	715	2	I40729	hypothetical 76k p
4	461.5	20.9	647	2	G74490	hypothetical prote
5	197	8.9	971	2	B90835	probable tail fibe
6	197	8.9	973	2	B85693	probable membrane
7	192.5	8.7	2055	2	T31110	extracellular matr
8	185.5	8.4	1822	2	S33441	EF protein - Strept
9	184	8.3	1122	2	G64887	probable tail fibe
10	181	8.2	2481	2	D90011	FmcB protein [impo
11	176.5	8.0	545	2	E84327	Htr7 transducer [i
12	176.5	8.0	545	2	T46811	halobacterial tran
13	175	7.9	627	2	F84194	Htr14 transducer [
14	173	7.8	2271	2	F90073	hypothetical prote
15	172.5	7.8	1128	2	T30296	R27-2 protein - Tr
16	168.5	7.6	2232	2	T34432	hypothetical prote
17	168	7.6	993	2	C90072	hypothetical prote
18	164	7.4	6713	2	B89321	hypothetical prote
19	163.5	7.4	2155	2	AD2742	conserved hypot
20	163.5	7.4	2155	2	C97523	hypothetical prote
21	160.5	7.3	544	2	T44938	transducer protein
22	160	7.3	536	1	A47190	transducer protein
23	160	7.3	536	2	E84318	Htr1 transducer [i
24	158.5	7.2	4776	2	E95206	cell wall surface
25	157	7.1	641	2	C82206	methyl-accepting c
26	156	7.1	810	2	T46810	halobacterial tran
27	156	7.1	810	2	F84327	Htr5 transducer [i
28	156	7.1	2186	2	H89960	hypothetical prote
29	154.5	7.0	1238	2	T03465	probable exonuclea

241 QTDATATQTEKGNALRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db
241 QTDATATQTEKGNALRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Qy 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVPNPGTTVGSGKQGGSSIGSIRVSMML 360
Db 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVPNPGTTVGSGKQGGSSIGSIRVSMML 360
Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSCAAQOELAAQAAKAGDDSAALADAQK 420
Db 361 DDAENETASILMSGFRQMIHMFNTENPDSCAAQOELAAQAAKAGDDSAALADAQK 420
Qy 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452
Db 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452

RESULT 2
E86581
CHLPN 76 kDa homolog 1 (CT622) [imported] - Chlamydomophila pneumoniae (strain J138)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: E86581
R:Shira, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-651 <S7C>
A:Cross-references: GB:BA000008; NID:g8979100; PIDN:BA98935.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: CPJ0728

Query Match 100.0%; Score 2204; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 7.2e-110;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTSVERWSI 60
Db 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTSVERWSI 60
Qy 61 LRSVNAALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTOAQATYDT 120
Db 61 LRSVNAALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTOAQATYDT 120
Qy 121 IFTSTSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKGAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKGAQITELAKY 180
Qy 181 ASDNQAILDSIGKLTSTFDLLQALLOSVAANNKAAELIKEMQDNVPVPGKTPALQAQSLVD 240
Db 181 ASDNQAILDSIGKLTSTFDLLQALLOSVAANNKAAELIKEMQDNVPVPGKTPALQAQSLVD 240
Qy 241 QTDATATQTEKGNALRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 241 QTDATATQTEKGNALRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Qy 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVPNPGTTVGSGKQGGSSIGSIRVSMML 360
Db 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVPNPGTTVGSGKQGGSSIGSIRVSMML 360
Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSCAAQOELAAQAAKAGDDSAALADAQK 420
Db 361 DDAENETASILMSGFRQMIHMFNTENPDSCAAQOELAAQAAKAGDDSAALADAQK 420
Qy 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452
Db 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452

RESULT 3
E86581
Hypothetical 76K protein - Chlamydomophila pneumoniae (strain AB39)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
C:Accession: I40729
R:Perez-Melgosa, M.; Kuo, C.
Infect. Immun. 62, 880-886, 1994
A:Title: Isolation and characterization of a gene encoding a Chlamydia pneumoniae 76-kil
A:Reference number: I40729; MUID:94156481; PMID:7509320
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-715 <RES>
A:Cross-references: GB:I23921; NID:g435961; PIDN:AA23117.1; PID:g435962
A:Experimental source: strain AR-39
C:Comment: This is the hypothetical translation of a sequence that was reported as two s

Query Match 99.9%; Score 2201; DB 2; Length 715;
Best Local Similarity 99.8%; Pred. No. 1.2e-109;
Matches 451; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTSVERWSI 60
Db 257 LVPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTSVERWSI 316
Qy 61 LRSVNAALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTOAQATYDT 120
Db 317 LRSVNAALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTOAQATYDT 376
Qy 121 IFTSTSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKGAQITELAKY 180
Db 377 IFTSTSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKGAQITELAKY 436
Qy 181 ASDNQAILDSIGKLTSTFDLLQALLOSVAANNKAAELIKEMQDNVPVPGKTPALQAQSLVD 240
Db 437 ASDNQAILDSIGKLTSTFDLLQALLOSVAANNKAAELIKEMQDNVPVPGKTPALQAQSLVD 496
Qy 241 QTDATATQTEKGNALRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 497 QTDATATQTEKGNALRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 556
Qy 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVPNPGTTVGSGKQGGSSIGSIRVSMML 360
Db 557 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVPNPGTTVGSGKQGGSSIGSIRVSMML 616
Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSCAAQOELAAQAAKAGDDSAALADAQK 420
Db 617 DDAENETASILMSGFRQMIHMFNTENPDSCAAQOELAAQAAKAGDDSAALADAQK 676
Qy 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452
Db 677 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 708

RESULT 4
G71490
Hypothetical protein CT622 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
N:Alternate names: chlpn 76kDa homolog CT622
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: G71490
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tr
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: G71490
A:Molecule type: DNA
A:Residues: 1-647 <ARN>
A:Cross-references: GB:AE001333; GB:AE001273; NID:g3329068; PIDN:AA068226.1; PID:g3329
A:Experimental source: serotype D, strain UW-3/Cx
C:Comment: This sequence was originally identified as homologous to part of a sequence
FIR:E72042).

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:04 ; Search time 560.656 Seconds
(without alignments)
11032.393 Million cell updates/sec

Title: US-10-608-559-5
Perfect score: 1456
Sequence: 1 ataaatctttaaaacag.....ctgttgagcgcaggagta 1456

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04.*

- 1: geneseq1980s.*
- 2: geneseq1990s.*
- 3: geneseq2000s.*
- 4: geneseq2001as.*
- 5: geneseq2001bs.*
- 6: geneseq2002s.*
- 7: geneseq2003as.*
- 8: geneseq2003bs.*
- 9: geneseq2003cs.*
- 10: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1456	100.0	1456	3 AAD02065	Aad02065 3'-trunca
2	1456	100.0	2238	3 AAD02066	Aad02066 C. pneumo
3	1455	99.9	2156	3 AAD02063	Aad02063 Chlamydia
4	1444	99.2	110000	2 AAX91990_08	Continuation (9 of
5	1355	93.1	1956	4 AAS57031	Aas57031 C. pneumo
6	1355	93.1	1956	6 ABL92612	Ab192612 Chlamydia
7	1355	93.1	1956	6 ABL91190	Ab191190 Chlamydia
8	1355	93.1	1956	9 ADD42825	Add42825 Chlamydia
9	1151	79.1	1852	3 AAD02064	Aad02064 5'-trunca
10	170	11.7	1550	3 AAX28411	Aax28411 Chlamydia
11	148.4	10.2	150	2 AAV16207	Aav16207 Part of t
12	55.4	3.8	2000	7 ADA71938	Ada71938 Rice gene
13	54.4	3.7	1941	6 ABL92619	Ab192619 Chlamydia
14	54.4	3.7	1944	9 ADD43867	Add43867 Chlamydia
15	51.8	3.6	110000	2 AAZ01425_07	Continuation (8 of
16	51.8	3.6	110000	2 AAZ01425_08	Continuation (9 of
17	51.6	3.5	2205	8 ADA28999	Ada28999 DNA encod
18	51.4	3.5	1537	3 AAG63292	Aag63292 C. tracho
19	51.4	3.5	1537	4 AAH56196	Aah56196 Chlamydia
20	51.4	3.5	1537	6 ABL92425	Ab192425 Chlamydia
21	49	3.4	896	2 AAQ68903	Aaq68903 PspA prot
22	49	3.4	946	2 AAX33124	Aax33124 Streptoco
23	49	3.4	957	2 AAT61726	Aat61726 Streptoco

24	49	3.4	1860	7 ACA49648	Aca49648 Prokaryot
25	49	3.4	1990	2 AAT61725	Aat61725 Streptoco
26	49	3.4	1990	7 ABX95373	Abx95373 S. pneumo
27	49	3.4	2085	2 AAQ78131	Aaq78131 Pneumococ
28	49	3.4	2085	2 AAT08979	Aat08979 DNA encod
29	49	3.4	2085	2 AAT07178	Aat07178 Pneumococ
30	49	3.4	2085	2 AAV33264	Aav33264 Pneumococ
31	49	3.4	2085	2 AAV39470	Aav39470 Streptoco
32	49	3.4	2085	2 AAZ25063	Aaz25063 Streptoco
33	49	3.4	2085	2 AAV84069	Aav84069 DNA encod
34	49	3.4	2085	2 AAX02012	Aax02012 S. pneumo
35	48.6	3.3	1171	4 AAS56981	Aas56981 C. tracho
36	48.6	3.3	1171	9 ADD42775	Add42775 Chlamydia
37	48.6	3.3	1834	4 AAS57006	Aas57006 C. tracho
38	48.6	3.3	1834	9 ADD42800	Add42800 Chlamydia
39	48.6	3.3	1983	4 AAS56996	Aas56996 C. tracho
40	48.6	3.3	1983	9 ADD42790	Add42790 Chlamydia
41	47.4	3.3	956	7 ABX95374	Abx95374 S. pneumo
42	47.4	3.3	2085	2 AAQ28674	Aaq28674 PspA gene
43	47.2	3.2	110000	6 ABA90521_07	Continuation (8 of
44	47	3.2	9373	6 ABL33291	Ab133291 Human imm
45	47	3.2	9373	6 ABK31317	Abk31317 Signal tr

ALIGNMENTS

RESULT 1
AAD02065
ID AAD02065 standard; DNA; 1456 BP.
XX
AC AAD02065;
XX
DT 26-MAR-2001 (first entry)
XX
DE 3'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.
XX
KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
KW upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; treatment; truncation mutant; ds.
XX
OS Chlamydophila pneumoniae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 101..1456
FT /*tag= a
FT /product= "3'-truncated Chlamydia pneumoniae 76KDa
FT /protein=
FT /note= "The coding region does not include stop codon"
FT /partial

WO2000066739-A2.
09-NOV-2000.
03-MAY-2000; 2000WO-CA000511.
03-MAY-1999; 99US-0132270P.
30-JUN-1999; 99US-0141276P.
(AVET) AVENTIS PASTEUR LTD.
Murdin AD, Oomen RP, Wang J, Dunn P;
WPI; 2000-687542/67.
P-PSDB; AAY71956.
Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
for vaccinating against Chlamydia infections.
Claim 2c; Page 102-104; 112pp; English.

Original signature

XX The present sequence is a DNA coding for 3'-truncated Chlamydia
CC pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the
CC diagnosis, prevention and treatment of C. pneumoniae infections (e.g.
CC pneumonia, upper respiratory tract disease, bronchitis, sinusitis and
CC acute respiratory disease such as cough, sore throat, hoarseness, fever;
CC and abnormal chest sounds on auscultation). C. pneumoniae sequence is
CC also used as vaccines for immunising humans against diseases caused by C.
CC pneumoniae
XX

SQ Sequence 1456 BP; 452 A; 308 C; 331 G; 365 T; 0 U; 0 Other;
Query Match 100.0%; Score 1456; DB 3; Length 1456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 60
DB 1 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 60
QY 61 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 120
DB 61 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 120
QY 121 AGGTCCTATAGACGAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 180
DB 121 AGGTCCTATAGACGAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 180
QY 181 GGGAGTCAGCAATAAGAGTCGCAAGCTCAAGCAATAGCAGGTCGGAAGCTTAAGCC 240
DB 181 GGGAGTCAGCAATAAGAGTCGCAAGCTCAAGCAATAGCAGGTCGGAAGCTTAAGCC 240
QY 241 TAAAGAACTTAAAGACGATTTCTAGAGCGATGAGGATCTTTCGCTTCTAGTGAATGC 300
DB 241 TAAAGAACTTAAAGACGATTTCTAGAGCGATGAGGATCTTTCGCTTCTAGTGAATGC 300
QY 301 TCTCATGAGTCGGCAGATAAGCTGGGATTTGCTTCTAGTGAATGCTTCTTCTACTAG 360
DB 301 TCTCATGAGTCGGCAGATAAGCTGGGATTTGCTTCTAGTGAATGCTTCTTCTACTAG 360
QY 361 CAGATCTGACGAGCTGACCTCAAGCAGCAGCGACCTACGCTTCTTCCACCCAGCTT 420
DB 361 CAGATCTGACGAGCTGACCTCAAGCAGCAGCGACCTACGCTTCTTCCACCCAGCTT 420
QY 421 TGATGATTAAGACTCAAGCGCAACAGCTTACGATCTTCTTCTTCTTCTTCTTCTTCT 480
DB 421 TGATGATTAAGACTCAAGCGCAACAGCTTACGATCTTCTTCTTCTTCTTCTTCTTCT 480
QY 481 AGCTGACATACAGGCTGCTTGGTGGAGCTTCCAGGATGCTGCTTCTTCTTCTTCTTCT 540
DB 481 AGCTGACATACAGGCTGCTTGGTGGAGCTTCCAGGATGCTGCTTCTTCTTCTTCTTCT 540
QY 541 AGCGGCTACTGATGAGAAACCGCAATCGCTGCGAGTGGAACTAAGATGCGCGATGC 600
DB 541 AGCGGCTACTGATGAGAAACCGCAATCGCTGCGAGTGGAACTAAGATGCGCGATGC 600
QY 601 AGTTAAAGTTGGCGCAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 660
DB 601 AGTTAAAGTTGGCGCAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 660
QY 661 TCTTGACTCTTTAGTAACTGACTTCTTCTGACCTTCTTCTGACCTTCTTCTTCTTCTTCT 720
DB 661 TCTTGACTCTTTAGTAACTGACTTCTTCTGACCTTCTTCTGACCTTCTTCTTCTTCTTCT 720
QY 721 TGTAGCAACAAATTAAGCAAGCTGAGCTTCTTAAAGAGATGCAAGATTAAGCAAGTGT 780
DB 721 TGTAGCAACAAATTAAGCAAGCTGAGCTTCTTAAAGAGATGCAAGATTAAGCAAGTGT 780
QY 781 CCCAGGAAACCGCTGCAATTTCTCAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
DB 781 CCCAGGAAACCGCTGCAATTTCTCAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
QY 841 ACAGATAGAAAGATGGAATGCGATTAGGATGCAATTTTTCGAGCAGACGAGCTAG 900

DB 841 ACAGATAGAAAGATGGAATGCGATTAGGATGCAATTTTTCGAGCAGACGAGCTAG 900
QY 901 TGGAGCTGTAGAAATCTAAATCTAATAACAGTATAGCAACATAGATTTCAGCTAAGC 960
DB 901 TGGAGCTGTAGAAATCTAAATCTAATAACAGTATAGCAACATAGATTTCAGCTAAGC 960
QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
DB 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
QY 1021 AATTCTTCAAGAGCGAAACAAATGGAATACAGGCTCAGAAAAGTCTTAAAAATATCAA 1080
DB 1021 AATTCTTCAAGAGCGAAACAAATGGAATACAGGCTCAGAAAAGTCTTAAAAATATCAA 1080
QY 1081 ACCTGCAGATGGTTCGTGATTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACA 1140
DB 1081 ACCTGCAGATGGTTCGTGATTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACA 1140
QY 1141 AGGAAGTAGTATTGGTAGTATTTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
DB 1141 AGGAAGTAGTATTGGTAGTATTTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
QY 1201 CGCTTCCATTTTGTGATGCTGGGTTTCGTGATGATTCACATGTTCAATAGCGAAATCC 1260
DB 1201 CGCTTCCATTTTGTGATGCTGGGTTTCGTGATGATTCACATGTTCAATAGCGAAATCC 1260
QY 1261 TGATTTCTCAAGTCGCCCAACAGGAGCTCGACACCAAGCTAGACGAGCGGAGCGCTGG 1320
DB 1261 TGATTTCTCAAGTCGCCCAACAGGAGCTCGACACCAAGCTAGACGAGCGGAGCGCTGG 1320
QY 1321 AGATGACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAGCGGCTTAGG 1380
DB 1321 AGATGACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAGCGGCTTAGG 1380
QY 1381 TAAAGCTGGGCAACCAACAGGCGCTGCGAGATGCTCAGAAAAGCTTTAGAGCGGCTTAGG 1440
DB 1381 TAAAGCTGGGCAACCAACAGGCGCTGCGAGATGCTCAGAAAAGCTTTAGAGCGGCTTAGG 1440
QY 1441 TGTGAGCGCAGAGTA 1456
DB 1441 TGTGAGCGCAGAGTA 1456

RESULT 2
AAD02066
ID AAD02066 standard; DNA; 2238 BP.
XX
AC AAD02066;
XX
DT 26-MAR-2001 (first entry)
XX
C. pneumoniae 76 kDa protein truncation mutant fusion gene.
76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
upper respiratory tract disease; bronchitis; sinusitis;
acute respiratory disease; cough; sore throat; hoarseness; fever;
vaccine; immunisation; treatment; truncation mutant; fusion gene; ds.
Chlamydia pneumoniae.
Synthetic.

Key Location/Qualifiers
misc_feature 1..665
/tag= a
/note= "This part of the sequence is unrelated to C.
pneumoniae 76 kDa gene"
CDS 766..2238
/tag= b
/product= "Truncated Chlamydia pneumoniae 76KDa protein"
misc_feature 2122..2238
/tag= c
/note= "This part of the sequence is unrelated to C."

FT XX pneumoniae 76 kDa gene"
PN WO200066739-A2.
XX
PD 09-NOV-2000.
XX
PF 03-MAY-2000; 2000WO-CA000511.
XX
PR 03-MAY-1999; 99US-0132270P.
XX
PR 30-JUN-1999; 99US-0141276P.
XX
XX (AVET) AVENTIS PASTEUR LTD.
XX
XX
XX Murdin AD, Oomen RP, Wang J, Dunn P;
PI
XX WPI; 2000-687542/67.
XX
DR P-PSDB; ANY71957.
XX
XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
PT for vaccinating against Chlamydia infections.
XX
XX Claim 32; Fig 3; 112pp; English.
XX
XX The present sequence is a DNA coding for a fusion protein comprising a
CC truncated Chlamydia pneumoniae 76 kDa protein and vector-encoded
CC residues. C. pneumoniae 76 kDa protein is used in the diagnosis,
CC prevention and treatment of C. pneumoniae infections (e.g. pneumonia,
CC upper respiratory tract disease, bronchitis, sinusitis and acute
CC respiratory disease such as cough, sore throat, hoarseness, fever; and
CC abnormal chest sounds on auscultation). C. pneumoniae sequence is also
CC used as vaccines for immunising humans against diseases caused by C.
XX pneumoniae
XX
SQ Sequence 2238 BP; 689 A; 459 C; 497 G; 593 T; 0 U; 0 Other;
Query Match 100.0%; Score 1456; DB 3; Length 2238;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATAAATCTTTAAACACAGGCTCGCATTAATATTAGTGAGAGCTTTTATTTTATTTT 60
DB 666 ATAAATCTTTAAACACAGGCTCGCATTAATATTAGTGAGAGCTTTTATTTTATTTT 725
QY 61 ATAAATCTTTAAACACAGGCTCGCATTAATATTAGTGAGAGCTTTTATTTTATTTT 120
DB 726 ATAAATCTTTAAACACAGGCTCGCATTAATATTAGTGAGAGCTTTTATTTTATTTT 785
QY 121 AGGTCTATAGCAAGCAAGACAGACACACCTCCGACAGATCTTTCTGCTCAAGGATGGA 180
DB 786 AGGTCTATAGCAAGCAAGACAGACACCTCCGACAGATCTTTCTGCTCAAGGATGGA 845
QY 181 GCGAGTGCAGCAATAAGAGTGCAGAGCTCAAGAAATAGAGGTGCGGAAGCTTAAGCC 240
DB 846 GCGAGTGCAGCAATAAGAGTGCAGAGCTCAAGAAATAGAGGTGCGGAAGCTTAAGCC 905
QY 241 TAAAGATCTAAGACCGATCTGTAGAGCATGAGAGCATCTTGGGTTCTGAGTGAATGC 300
DB 906 TAAAGATCTAAGACCGATCTGTAGAGCATGAGAGCATCTTGGGTTCTGAGTGAATGC 965
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAAACAGCTGCTTCTTACTAG 360
DB 966 TCTCATGAGTCTGGCAGATAAGCTGGGTTATTGCTTCTAGTAAACAGCTGCTTCTTACTAG 1025
QY 361 CAGATCTGCAGACGTGGAATCAACACAGCAGCGCCGACCTAGCCCTCTCCACCCACGTT 420
DB 1026 CAGATCTGCAGACGTGGAATCAACACAGCAGCGCCGACCTAGCCCTCTCCACCCACGTT 1085
QY 421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCACT 480
DB 1086 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCACT 1145
QY 481 AGCTGACATACAGGCTGCTTGGTGGAGCTCCAGGATGCTGTCACCTAATATAAGGATAC 540

DB 1146 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGTCACTAATATAAGGATAC 1205
QY 541 AGCGGCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAAACTTAAGAAATGCCGATGC 600
DB 1206 AGCGGCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAAACTTAAGAAATGCCGATGC 1265
QY 601 AGTTAAAGTTGGCGCCCAAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGAT 660
DB 1266 AGTTAAAGTTGGCGCCCAAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGAT 1325
QY 661 TCTTGACTCTTTAGGTAAACTGACTTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC 720
DB 1326 TCTTGACTCTTTAGGTAAACTGACTTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC 1385
QY 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATCGAAGATTAACCCAGTAGT 780
DB 1386 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATCGAAGATTAACCCAGTAGT 1445
QY 781 CCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
DB 1446 CCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 1505
QY 841 ACAGATAGAGAAAGATGGAATCGGATTAGGATGCAATTTTTCAGGACAGAAACGCTAG 900
DB 1506 ACAGATAGAGAAAGATGGAATCGGATTAGGATGCAATTTTTCAGGACAGAAACGCTAG 1565
QY 901 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTTCAGCTAAGC 960
DB 1566 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTTCAGCTAAGC 1625
QY 961 AGCAATCGCTACTCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCC 1020
DB 1626 AGCAATCGCTACTCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCC 1685
QY 1021 AATTCTTCAAGACGGAACAAATGGTAATACAGGCTGAGAAAAATCTTAAAAATATCAA 1080
DB 1686 AATTCTTCAAGACGGAACAAATGGTAATACAGGCTGAGAAAAATCTTAAAAATATCAA 1745
QY 1081 ACCTGCGAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGCAACA 1140
DB 1746 ACCTGCGAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGCAACA 1805
QY 1141 AGGAAGTAGTATTGCTAGTATTCTGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 1200
DB 1806 AGGAAGTAGTATTGCTAGTATTCTGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 1865
QY 1201 CGCTTCCATTTTGTGTTCTGGTTTCGTCAGATGATTCATGTTCAATACGAAAAATCC 1260
DB 1866 CGCTTCCATTTTGTGTTCTGGTTTCGTCAGATGATTCATGTTCAATACGAAAAATCC 1925
QY 1261 TGATTTCAAGCTGCCAAGCAGGAGCTGCGAGCAACAGCTAGAGCAGGAAAGCGGCTGG 1320
DB 1926 TGATTTCAAGCTGCCAAGCAGGAGCTGCGAGCAACAGCTAGAGCAGGAAAGCGGCTGG 1985
QY 1321 AGATTGACAGTGTGCTGCGAGGCTGGCAGATGCTCAGAAAGCTTTTGAAGCGGCTCTAGG 1380
DB 1986 AGATTGACAGTGTGCTGCGAGGCTGGCAGATGCTCAGAAAGCTTTTGAAGCGGCTCTAGG 2045
QY 1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTTAGGACAGATCGGTTCTGCTGCTGT 1440
DB 2046 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTTAGGACAGATCGGTTCTGCTGCTGT 2105
QY 1441 TGTAGCGCAGAGTA 1456
DB 2106 TGTAGCGCAGAGTA 2121

RESULT 3
AAD02063
ID AAD02063 standard; DNA; 2156 BP.
XX
AC AAD02063;
XX

DT	15-SEP-2003 (revised)	QY	301	TCTCATGAGTCTGGCAGATAAAGCTGGGTATTGCTTCTAGTAAACAGCTCGCTCTTACTAG	360
XX	26-MAR-2001 (first entry)	DB	301	TCTCATGAGTCTGGCAGATAAAGCTGGGTATTGCTTCTAGTAAACAGCTCGCTCTTACTAG	360
DE	Chlamydia pneumoniae 76 kDa full-length protein encoding DNA.	QY	361	CAGATCTGCAGAGCTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCAGCTT	420
XX	76 kDa protein; bactericidal; diagnosis; prevention; treatment;	DB	361	CAGATCTGCAGAGCTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCAGCTT	420
KW	pneumonia; upper respiratory tract disease; bronchitis; sinusitis;	QY	421	TGATGATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTACCTCAACATCACT	480
KW	acute respiratory disease; cough; sore throat; hoarseness; fever;	DB	421	TGATGATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTACCTCAACATCACT	480
KW	vaccine; immunisation; ds.	QY	481	AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGCTCACTAATAAAGGATAC	540
XX	Chlamydia pneumoniae.	DB	481	AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGCTCACTAATAAAGGATAC	540
OS	Key Location/Qualifiers	QY	541	AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC	600
XX	101..2056	DB	541	AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC	600
XX	/*tag= a	QY	601	AGTTAAAGTTGCGCGCAAAATTAAGAAATAGCGAATATGCTTCGGATACCAAGCGAT	660
XX	/product= "Chlamydia pneumoniae 76kDa protein"	DB	601	AGTTAAAGTTGCGCGCAAAATTAAGAAATAGCGAATATGCTTCGGATACCAAGCGAT	660
XX	WC200066739-A2.	QY	661	TCTTGACTCTTTAGTAACTGACTTCTCCTCGACCTCTTACAGGCTGCTCTTCTCAATC	720
XX	09-NOV-2000.	DB	661	TCTTGACTCTTTAGTAACTGACTTCTCCTTGAACCTTTACAGGCTGCTCTTCTCAATC	720
XX	03-MAY-2000; 2000WC-CA000511.	QY	721	TGTAGCAAAACATAA CAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	780
XX	03-MAY-1999; 99US-0132270P.	DB	721	TGTAGCAAAACATAA CAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	780
XX	30-JUN-1999; 99US-0141276P.	QY	781	CCAGGGAAGAACGCTGCAATGCTCAATCTTTAGTTCAGACAGATGCTTACAGCGAC	840
XX	(AVER) AVENTIS PASTEUR LTD.	DB	781	CCAGGGAAGAACGCTGCAATGCTCAATCTTTAGTTCAGACAGATGCTTACAGCGAC	840
XX	Murdin AD, Omen RP, Wang J, Dunn P;	QY	841	ACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATATTTTGCAGGACAGAAACGCTAG	900
XX	WPI; 2000-687542/67.	DB	841	ACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATATTTTGCAGGACAGAAACGCTAG	900
XX	P-PSDB; AAY71954.	QY	901	TGGAGCTGTAGAAAATGCTTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGC	960
XX	Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful	DB	901	TGGAGCTGTAGAAAATGCTTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGC	960
XX	for vaccinating against Chlamydia infections.	QY	961	AGCAATCGCTACTGTCTAGACACAAATAGCTCAAGCTCAGAAAAGTTCCCGACTCTCC	1020
XX	Claim 2a; Fig 1; 112pp; English.	DB	961	AGCAATCGCTACTGTCTAGACACAAATAGCTCAAGCTCAGAAAAGTTCCCGACTCTCC	1020
XX	The present sequence is a DNA coding for Chlamydia pneumoniae 76 kDa	QY	1021	AATTTCTTCAAGAACGCGAAACAAATGGTAATACAGGCTGAGAAAGATCTTTAAATAATCAA	1080
XX	protein. C. pneumoniae 76 kDa protein is used in the diagnosis,	DB	1021	AATTTCTTCAAGAACGCGAAACAAATGGTAATACAGGCTGAGAAAGATCTTTAAATAATCAA	1080
XX	prevention and treatment of C. pneumoniae infections (e.g. pneumonia,	QY	1081	ACCTGAGATGTTCTGATGTTCCAAATCCAGGAACATACAGTTGGAGGCTCCCAAGCAACA	1140
XX	upper respiratory tract disease, bronchitis, sinusitis and acute	DB	1081	ACCTGAGATGTTCTGATGTTCCAAATCCAGGAACATACAGTTGGAGGCTCCCAAGCAACA	1140
XX	respiratory disease such as cough, sore throat, hoarseness, fever, and	QY	1141	AGGAAGTAGTATTGGTAGTATTTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC	1200
XX	abnormal chest sounds on auscultation). C. pneumoniae sequence is also	DB	1141	AGGAAGTAGTATTGGTAGTATTTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC	1200
XX	used as vaccines for immunising humans against diseases caused by C.	QY	1201	CGTTTCCATTTTGTATGCTGCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATAATCC	1260
XX	pneumoniae. (Updated on 15-SEP-2003 to standardise OS field)	DB	1201	CGTTTCCATTTTGTATGCTGCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATAATCC	1260
XX	Sequence 2156 BP; 672 A; 461 C; 471 G; 552 T; 0 U; 0 Other;	QY	1261	TGATTCTCAAGCTGCCCAACAGGAGCTCCGAGCAACAGCTAGACGAGCGAAAGCCGCTGG	1320
XX	Query Match 99.9%; Score 1455; DB 3; Length 2156;	DB	1261	TGATTCTCAAGCTGCCCAACAGGAGCTCCGAGCAACAGCTAGACGAGCGAAAGCCGCTGG	1320
XX	Best Local Similarity 100.0%; Pred.No. 0;	QY	1321	AGATGACAGTGTCTGCTGCGCTGCAGATGCTCAGAAAAGCTTTTGAAGCGGCTCTTAGG	1380
XX	Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	DB	1321	AGATGACAGTGTCTGCTGCGCTGCAGATGCTCAGAAAAGCTTTTGAAGCGGCTCTTAGG	1380
XX	1 ATAAATCTTAAACAGAGCTCGCATTAATTTAGTGAGAGCTTTTATTTTATTTT 60	QY	1381	TAAAGCTGGGCAACAAACAGGSCATACCTCAATGCTTTTAGGACAGATCGCTTCTGCTGT	1440
XX	1 ATAAATCTTAAACAGAGCTCGCATTAATTTAGTGAGAGCTTTTATTTTATTTT 60	DB			
XX	61 ATAAATCTTAAACAGAGCTCGCATTAATTTAGTGAGAGCTTTTATTTTATTTT 120	QY			
XX	61 ATAAATCTTAAACAGAGCTCGCATTAATTTAGTGAGAGCTTTTATTTTATTTT 120	DB			
XX	121 AGGTCTTATAGCAAGAACAGACGACACCTCCCGCAGATCTTTCTGCTCAAGGATTTGA 180	QY			
XX	121 AGGTCTTATAGCAAGAACAGACGACACCTCCCGCAGATCTTTCTGCTCAAGGATTTGA 180	DB			
XX	181 GCGAGTGCACCAATAGAGTGGGAGCTCAAGATAGCAGTGGGAGCTTAAGCC 240	QY			
XX	181 GCGAGTGCACCAATAGAGTGGGAGCTCAAGATAGCAGTGGGAGCTTAAGCC 240	DB			
XX	241 TAAAGATCTAAGACCGATTCTGTAGAGCGATGAGCATCTTGGCTTCTGAGTGAATGC 300	QY			
XX	241 TAAAGATCTAAGACCGATTCTGTAGAGCGATGAGCATCTTGGCTTCTGAGTGAATGC 300	DB			

[Handwritten signature]

Db	1381	TAAAGCTGGGCAACAACAGGGCACTCAATGCTTTAGGACAGATCGCTCTCTGCTGCT	1440
Qy	1441	TGTGAGCGCAGGAGT	1455
Db	1441	TGTGAGCGCAGGAGT	1455
RESULT 4			
AAx91990_08/c			
Continuation (9 of 13) of AAx91990 from base 800001 (Nucleotide sequence of the complete			
WP Sequence split into 13 fragments LOCUS AAx91990 Accession AAx91990			
WP	Fragment Name	Begin	End
WP	AAx91990_00	1	110000
WP	AAx91990_01	100001	210000
WP	AAx91990_02	200001	310000
WP	AAx91990_03	300001	410000
WP	AAx91990_04	400001	510000
WP	AAx91990_05	500001	610000
WP	AAx91990_06	600001	710000
WP	AAx91990_07	700001	810000
WP	AAx91990_08	800001	910000
WP	AAx91990_09	900001	1010000
WP	AAx91990_10	1000001	1110000
WP	AAx91990_11	1100001	1210000
WP	AAx91990_12	1200001	12300025
Query Match 99.2%; Score 1444; DB 2; Length 110000;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			
Qy	1	ATAAAATCTTTAAACACGGCTCGCATTAATTTAGTGAGAGCTTTTTTTTTTTTTTTT	60
Db	28698	ATAAAATCTTTAAACACGGCTCGCATTAATTTAGTGAGAGCTTTTTTTTTTTTTTTT	28639
Qy	61	ATAATAAACTAAAGATTTTATTATTTTATTTTATGTTTATGTTTAACTATTTGGTCC	120
Db	28638	ATAATAAACTAAAGATTTTATTATTTTATGTTTATGTTTAACTATTTGGTCC	28579
Qy	121	AGGTCTATAGACGAAACAGAACGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA	180
Db	28578	AGGTCTATAGACGAAACAGAACGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA	28519
Qy	181	GGCGAGTCAGCAATATAGAGTCGGGAAGCTCAAGAAATAGCAGTGGGAGCTAAGCC	240
Db	28518	GGCGAGTCAGCAATATAGAGTCGGGAAGCTCAAGAAATAGCAGTGGGAGCTAAGCC	28459
Qy	241	TAAAGAACTTAAGACCGATTCTGTAGCGATGGAGCATCTTGCCTTCTGCAAGTGAATGC	300
Db	28458	TAAAGAACTTAAGACCGATTCTGTAGAGCGGATGGAGCATCTTGCCTTCTGCAAGTGAATGC	28399
Qy	301	TCTCATGAGTCGGCAGATAGCTGGGATTTGCTTTCTAGTAACAGCTCGTCTTCTACTAG	360
Db	28398	TCTCATGAGTCGGCAGATAGCTGGGATTTGCTTTCTAGTAACAGCTCGTCTTCTACTAG	28339
Qy	361	CAGATCTGAGCTGGATCTCAACACAGCGACCGACCTAGCCCTCCACCCACAGTT	420
Db	28338	CAGATCTGAGCTGGATCTCAACACAGCGACCGACCTAGCCCTCCACCCACAGTT	28279
Qy	421	TGATGATTATAGATCTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT	480
Db	28278	TGATGATTATAGATCTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT	28219
Qy	481	AGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGCTCACTAATATAAGGATAC	540
Db	28218	AGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGCTCACTAATATAAGGATAC	28159
Qy	541	AGCGGCTACTGATAGGAAACCGCAATCGCTGGGAGTGGAACTTAAGATGCGGATGC	600
Db	28158	AGCGGCTACTGATAGGAAACCGCAATCGCTGGGAGTGGAACTTAAGATGCGGATGC	28099
Qy	601	AGTTAAAGTTGGCGCAAAATTACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT	660

RESULT 5

AAS57031

ID AAS57031 standard; DNA; 1956 BP.

XX

AC AAS57031;

XX

11-SEP-2003 (revised)

DT 16-JAN-2002 (first entry)

XX

C. pneumoniae DNA encoding the Ctr622 homologue CPn0728.

DE

Chlamydia; ds; sexually transmitted disease; PID; antibacterial;

KW pelvic inflammatory disease; antigen; trachoma; gynecological;

KW acute respiratory tract infection; atherosclerosis; male infertility;

KW	coronary heart disease.	QY	521	GTCACTAATAAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG	580
XX		Db	421	GTCACTAATAAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG	480
OS	Chlamydia pneumoniae.	QY	581	GAATCTAAGAAATGCGATGCGATTAAGTTGCGCGGCAAAATTAAGAAATAGCGAAATAT	640
XX		Db	481	GAATCTAAGAAATGCGATGCGATTAAGTTGCGCGGCAAAATTAAGAAATAGCGAAATAT	540
FN	WO200181379-A2.	QY	641	GCTTCGGATACCAAGCGATTTCTGACTCTTTAGTAACTGACTCTCTCTCGACTCTCTTA	700
PD	01-NOV-2001.	Db	541	GCTTCGGATACCAAGCGATTTCTGACTCTTTAGTAACTGACTCTCTCTCGACTCTCTTA	600
XX	23-APR-2001; 2001WO-US013081.	QY	701	CAGGTGCTCTTCTTCCAAATCTGTAGCAAAACAATAAAGAGAGTGTCTTTAAAGAG	760
XX	21-APR-2000; 2000US-0198853P.	Db	601	CAGGTGCTCTTCTTCCAAATCTGTAGCAAAACAATAAAGAGAGTGTCTTTAAAGAG	660
PR	20-JUL-2000; 2000US-0219752P.	QY	761	ATGCAAGATTAACCCAGTAGTCTCCAGGAAAGCGCTGCAATTCCTTCTTCTTCTTCTT	820
XX	(CORI-) CORIXA CORP.	Db	661	ATGCAAGATTAACCCAGTAGTCTCCAGGAAAGCGCTGCAATTCCTTCTTCTTCTTCTT	720
PA		QY	821	CAGCAGATGCTTACAGCGACACAGATAGAGAAAGATGCAATTCGATTTAGGATGCATAT	880
PI	Bhatia A, Probst P, Stromberg EJ;	Db	721	CAGCAGATGCTTACAGCGACACAGATAGAGAAAGATGCAATTCGATTTAGGATGCATAT	780
XX	WPI; 2001-616771/71.	QY	881	TTTGAGGACAGACGCTAGTGGAGCTGTAGAAATGTAAATCTTAATAACAGTATACG	940
DR	P-PSDB; RAU38899.	Db	781	TTTGAGGACAGACGCTAGTGGAGCTGTAGAAATGTAAATCTTAATAACAGTATACG	840
XX	New polynucleotide for treating Chlamydia infections encodes a	QY	941	AACATAGATTCAGCTTAAGCAGCAATCGCTACTGTCTAGACACAATACTGCTGAGCTCAG	1000
XX	polynucleotides containing an immunogenic portion of a Chlamydia antigen.	Db	841	AACATAGATTCAGCTTAAGCAGCAATCGCTACTGTCTAGACACAATACTGCTGAGCTCAG	900
PS	Disclosure; Page 143; 208pp; English.	QY	1001	AAAAAGTTCCCGGACTCTCCAAATTCCTTCAAGAGCGGAAACAAATGCTAAATCAGGCTGAG	1060
CC	The invention relates to isolated polynucleotide encoding at least a	Db	901	AAAAAGTTCCCGGACTCTCCAAATTCCTTCAAGAGCGGAAACAAATGCTAAATCAGGCTGAG	960
CC	partial Chlamydia protein which is an antigenic fragment, or the	QY	1061	AAAGATCTTAAATAATCAAACTGAGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT	1120
CC	complements, fragments, homologues and variants, and antibodies raised	Db	961	AAAGATCTTAAATAATCAAACTGAGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT	1020
CC	against the antigenic proteins (or fragments). The nucleic acids,	QY	1121	GTGAGGCTCCAGCAACAGGAAAGTAGTATTGTTAGTATTGTTAGTATTGTTAGTATTGTTAGT	1180
CC	proteins and antibodies are used to diagnose and treat Chlamydia	Db	1021	GTGAGGCTCCAGCAACAGGAAAGTAGTATTGTTAGTATTGTTAGTATTGTTAGTATTGTTAGT	1080
CC	infections (e.g. a sexually transmitted disease, pelvic inflammatory	QY	1181	GATGATGCTGAAATGAGACCGCTTCCATTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT	1240
CC	disease (PID), acute respiratory tract infection, trachoma, and in the	Db	1081	GATGATGCTGAAATGAGACCGCTTCCATTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
CC	atherosclerosis and coronary heart disease). The compounds of the invention are also	QY	1241	ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGGAGCTCCGAGCACAAGCT	1300
CC	treatment of male infertility. The compounds of the invention are also	Db	1141	ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGGAGCTCCGAGCACAAGCT	1200
CC	useful for detecting the presence of Chlamydia in a patient, and	QY	1301	AGAGCAGCGAAAGCGCTGGAGATGACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1360
CC	stimulating and/or expanding T cells specific for a Chlamydia protein.	Db	1201	AGAGCAGCGAAAGCGCTGGAGATGACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260
CC	The present sequence encodes a Chlamydia antigen. (Updated on 11-SEP-2003	QY	1361	GCTTTAGAACGCGCTCTAGGTAAAGCTGGGCAACAACAGGCGCATCTCAATGCTTTAGGA	1420
CC	to standardise OS field)	Db	1261	GCTTTAGAACGCGCTCTAGGTAAAGCTGGGCAACAACAGGCGCATCTCAATGCTTTAGGA	1320
XX		QY	1421	CAGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1480
XX		Db	1321	CAGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1380
XX		QY	1481	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1540
XX		Db	1381	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1340
XX		QY	1541	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1600
XX		Db	1341	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1300
XX		QY	1601	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1660
XX		Db	1301	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260
XX		QY	1661	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1720
XX		Db	1261	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1220
XX		QY	1721	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1780
XX		Db	1221	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1180
XX		QY	1781	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1840
XX		Db	1181	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
XX		QY	1841	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1900
XX		Db	1141	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1100
XX		QY	1901	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1960
XX		Db	1101	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1060
XX		QY	1961	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2020
XX		Db	1061	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1020
XX		QY	2021	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2080
XX		Db	1021	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	980
XX		QY	2081	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2140
XX		Db	981	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	940
XX		QY	2141	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2200
XX		Db	941	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900
XX		QY	2201	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2260
XX		Db	901	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2220
XX		QY	2261	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2320
XX		Db	2221	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2180
XX		QY	2321	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2380
XX		Db	2181	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2140
XX		QY	2381	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2440
XX		Db	2141	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2100
XX		QY	2441	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2500
XX		Db	2101	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2060
XX		QY	2501	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2560
XX		Db	2061	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2020
XX		QY	2561	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2620
XX		Db	2021	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2580
XX		QY	2621	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2680
XX		Db	2581	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2640
XX		QY	2681	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2740
XX		Db	2641	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2700
XX		QY	2741	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2800
XX		Db	2701	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2760
XX		QY	2801	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2860
XX		Db	2761	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2820
XX		QY	2861	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2920
XX		Db	2821	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2880
XX		QY	2921	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3020
XX		Db	2881	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2980
XX		QY	3021	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3080
XX		Db	2981	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3040
XX		QY	3081	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3140
XX		Db	3041	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3100
XX		QY	3141	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3200
XX		Db	3101	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3160
XX		QY	3201	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3260
XX		Db	3161	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3220
XX		QY	3261	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3320
XX		Db	3221	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3280
XX		QY	3321	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3420
XX		Db	3281	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3380
XX		QY	3421	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3520
XX		Db	3381	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3480
XX		QY	3521	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3620
XX		Db	3481	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3580
XX		QY	3621	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3720
XX		Db	3581	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3680
XX		QY	3721	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3820
XX		Db	3681	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3780
XX		QY	3821	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3920
XX		Db	3781	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3880
XX		QY	3921	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4020
XX		Db	3881	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3980
XX		QY	4021	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4120
XX		Db	3981	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4080
XX		QY	4121	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4220
XX		Db	4081	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4180
XX		QY	4221	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4320
XX		Db	4181	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4280
XX		QY	4321	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4420
XX		Db	4281	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4380
XX		QY	4421	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4520
XX		Db	4381	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4480
XX		QY	4521	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4620
XX		Db	4481	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4580
XX		QY	4621	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4720
XX		Db	4581	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4680
XX		QY	4721	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4820
XX		Db	4681	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4780
XX		QY	4821	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4920
XX		Db	4781	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4880
XX		QY	4921	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5020
XX		Db	4881	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4980
XX		QY	5021	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5120
XX		Db	4981	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5080
XX		QY	5121	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5220
XX		Db	5081	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5180
XX		QY	5221	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5320
XX		Db	5181	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5280
XX		QY	5321	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5420
XX		Db	5281	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5380
XX		QY	5421	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5520
XX		Db	5381	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5480
XX		QY	5521	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5620
XX		Db	5481	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5580
XX		QY	5621	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5720
XX		Db	5581	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5680
XX		QY	5721	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5820
XX		Db	5681	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5780
XX		QY	5821	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5920
XX		Db	5781	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5880
XX		QY	5921	AGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6020
XX		Db	5881		

AC	ABL91190;	DB	1	ATGGTTAATCTATTTGGTCCAGGTCCTATATACGAAACAGAACGCCACCTCCCGCAGAT	60
XX					
XX	29-AUG-2003 (revised)	QY	161	CTTTCTGCTCAAGGANTGGAGGCGAGTCAGCAATAAGAGCTGGGGAAGCTCAAGAATA	220
DT	29-JUL-2002 (first entry)				
XX	Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.	DB	61	CTTTCTGCTCAAGGANTGGAGGCGAGTCAGCAATAAGAGCTGGGGAAGCTCAAGAATA	120
DE					
XX	Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.	QY	221	GCAGGTGGGAAGCTAAGCCCTAAAGATCTAAGACCGATTCTGTAGACGATGGAGCATC	280
XX	Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.	DB	121	GCAGGTGGGAAGCTAAGCCCTAAAGATCTAAGACCGATTCTGTAGACGATGGAGCATC	180
KW	human respiratory disease; cardiovascular disease; atherosclerosis;	QY	281	TTGCGTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGT	340
KW	coronary artery disease; carotid artery stenosis; myocardial infarction;	DB	181	TTGCGTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGT	240
KW	cerebrovascular disease; aortic aneurysm; claudication; stroke;	QY	341	AAACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGATCTCAACGACGACGCCACCT	400
XX	strain CWL029; open reading frame; ORF; gene; ds.	DB	241	AAACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGATCTCAACGACGACGCCACCT	300
XX	Chlamydia pneumoniae.	QY	401	ACGCTCTCTCCACCCAGCTTTGATGATTTATAGACTCAAGCGCAACAGCTTACGATCT	460
XX		DB	301	ACGCTCTCTCCACCCAGCTTTGATGATTTATAGACTCAAGCGCAACAGCTTACGATCT	360
XX		QY	461	ATCTTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTTGTGAGCTCCAGGATGCT	520
XX		DB	361	ATCTTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTTGTGAGCTCCAGGATGCT	420
XX		QY	521	GTCACTAATATAAAGGATACAGCGCTACTGATAGGAGAAACCGCAATGCTCGGAGTGG	580
XX		DB	421	GTCACTAATATAAAGGATACAGCGCTACTGATAGGAGAAACCGCAATGCTCGGAGTGG	480
XX		QY	581	GAACTAAGAAATGCCGATGCTTAAGTTGGCGCGCAATTTACAGAAATTACGAAATAT	640
XX		DB	481	GAACTAAGAAATGCCGATGCTTAAGTTGGCGCGCAATTTACAGAAATTACGAAATAT	540
XX		QY	641	GCTTCGGATAACCAAGCGATTTCTGACTCTTTAGGTAAACTGACTCTCTTCGACCTCTTA	700
XX		DB	541	GCTTCGGATAACCAAGCGATTTCTGACTCTTTAGGTAAACTGACTCTCTTCGACCTCTTA	600
XX		QY	701	CAGGCTGCTCTTCTCCAAATCTGTAGCAACAATAAAGAGAGCTGAGCTTCTTAAAGAG	760
XX		DB	601	CAGGCTGCTCTTCTCCAAATCTGTAGCAACAATAAAGAGAGCTGAGCTTCTTAAAGAG	660
XX		QY	761	ATGCAAGATAACCCAGTACTGCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGAT	820
XX		DB	661	ATGCAAGATAACCCAGTACTGCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGAT	720
XX		QY	821	CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAAATCGGATTAGGATGCAAT	880
XX		DB	721	CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAAATCGGATTAGGATGCAAT	780
XX		QY	881	TTTTCGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGC	940
XX		DB	781	TTTTCGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGC	840
XX		QY	941	AAACATAGATTACAGTAAAGCAGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTAG	1000
XX		DB	841	AAACATAGATTACAGTAAAGCAGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTAG	900
XX		QY	1001	AAAAAGTTCGCCGACTCTCCAAATCTTCAAGAACGCGAAACAAATGTTAATACAGGCTGAG	1060
XX		DB	901	AAAAAGTTCGCCGACTCTCCAAATCTTCAAGAACGCGAAACAAATGTTAATACAGGCTGAG	960
XX		QY	1061	AAAGATCTTAAATAATCAAACTGAGATGTTCTGATGTTCCCAATCCAGGAATCTACA	1120
XX		DB	961	AAAGATCTTAAATAATCAAACTGAGATGTTCTGATGTTCCCAATCCAGGAATCTACA	1020
XX		QY	1121	GTTCGAGGCTCCAGCAACAAGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTTTA	1180
XX		DB	1021	GTTCGAGGCTCCAGCAACAAGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTTTA	1080
XX		QY	1181	GATGATGCTGAAATGAGACCGCTTCCATTTTGTATGCTCTGGGTTTCGTGATGATTCAC	1240

ABL91190; 29-AUG-2003 (revised) 29-JUL-2002 (first entry) Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14. Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14. Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14. human respiratory disease; cardiovascular disease; atherosclerosis; coronary artery disease; carotid artery stenosis; myocardial infarction; cerebrovascular disease; aortic aneurysm; claudication; stroke; strain CWL029; open reading frame; ORF; gene; ds. Chlamydia pneumoniae. Key Location/Qualifiers CDS 1..1956 /tag= a /product= "cp7033" W0200202606-A2. 10-JAN-2002. 03-JUL-2001; 2001W0-IB001445. 03-JUL-2000; 2000GB-00016363. 14-JUL-2000; 2000GB-00017047. 26-JUL-2000; 2000GB-00017983. 07-AUG-2000; 2000GB-00019368. 18-AUG-2000; 2000GB-00020440. 14-SEP-2000; 2000GB-00022583. 10-NOV-2000; 2000GB-00027549. 22-DEC-2000; 2000GB-00031706. (CHIR-) CHIRON SPA. Ratti G, Grandi G; WPI; 2002-154726/20. N-PSDB; ABB90532. Novel Chlamydia pneumoniae protein useful in the manufacture of a medicament for treatment or prevention of infection due to Chlamydia, preferably Chlamydia pneumoniae, and for diagnostic purposes. Claim 5; Page 47-48; 364pp; English. Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding them. The proteins are predicted to be immunogenic and may therefore be useful in vaccine production and for diagnostic purposes. Chlamydia pneumoniae is a common cause of respiratory disease in humans, and is also involved in the development of cardiovascular diseases such as atherosclerosis, coronary artery disease, carotid artery stenosis, myocardial infarction, cerebrovascular disease, aortic aneurysm, claudication and stroke. The proteins and nucleic acids of the invention may be used in vaccines and pharmaceutical compositions for the prevention or treatment of chlamydial infections, particularly Chlamydia pneumoniae infections. The proteins may also be used in the detection of Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched DNA probe assay or blotting techniques for determining Chlamydia pneumoniae gene expression. The present sequence represents a specifically claimed DNA which encodes a Chlamydia pneumoniae protein of the invention. (Updated on 29-AUG-2003 to standardise OS field) Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other; Query Match 93.1%; Score 1355; DB 6; Length 1956; Best Local Similarity 100.0%; Pred. No. 0; Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0; 101 ATGGTTAATCTATTTGGTCCAGGTCCTATATACGAAACAGAACGCCACCTCCCGCAGAT 160

QY 1085 GCAGATGGTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAGGA 1144
DB 781 GCAGATGGTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAGGA 840
QY 1145 AGTAGTATTGGTAGTATTCGTGTTTCATGCTCTTATGATGATGCTGAAATGAGCGCT 1204
DB 841 AGTAGTATTGGTAGTATTCGTGTTTCATGCTCTTATGATGATGCTGAAATGAGCGCT 900
QY 1205 TCCATTTTATGATGCTGGGTTTCGTGATGATGATTCACATGTTCAATACGGAAATCCTGAT 1264
DB 901 TCCATTTTATGATGCTGGGTTTCGTGATGATGATTCACATGTTCAATACGGAAATCCTGAT 960
QY 1265 TCTCAAGTCCCAACAGGAGCTCGGACACAGCTAGACAGCAAGCGCTGGAGAT 1324
DB 961 TCTCAAGTCCCAACAGGAGCTCGGACACAGCTAGACAGCAAGCGCTGGAGAT 1020
QY 1325 GACAGTCTGCTCAGCGCTGGAGATGCTCAGAAAGCTTTAGAGCGCTCTAGGTAA 1384
DB 1021 GACAGTCTGCTCAGCGCTGGAGATGCTCAGAAAGCTTTAGAGCGCTCTAGGTAA 1080
QY 1385 GCTGGGCAACACAGGAGCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTG 1444
DB 1081 GCTGGGCAACACAGGAGCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTG 1140
QY 1445 AGCGCAGGAGT 1455
DB 1141 AGCGCAGGAGT 1151

RESULT 10
ID AAA28411 standard; DNA; 1550 BP.
XX
AC AAA28411;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Chlamydia pneumoniae lorf2 coding sequence.
XX
KW Lorf2; vaccine; antibacterial; antigen; ss.
XX
OS Chlamydia pneumoniae.

XX Key Location/Qualifiers
XX CDS 101..1369
XX FT /*tag= a
XX
XX WO200024901-A1.
XX
XX PD 04-MAY-2000.
XX
XX PF 28-OCT-1999; 99WO-GB003565.
XX
XX PR 28-OCT-1998; 98US-0106037P.
XX PR 20-SEP-1999; 99US-0154658P.
XX PR 26-OCT-1999; 99US-00427501.
XX
XX PA (CONN-) CONNAUGHT LAB LTD.
XX
XX PI Murdin AD, Oomen RP, Dunn PL;
XX
XX DR WPI; 2000-350742/30.
XX
XX DR P-PSDB; AAY92716.
XX

PT Isolated polynucleotide encoding a Chlamydia polypeptide useful to treat,
PT diagnose and prevent disease caused by Chlamydia infection.
XX
XX Claim 1; Fig 1A-C; 88pp; English.
XX
XX This sequence encodes lorf2 protein of a strain of Chlamydia pneumoniae.
XX
XX Comparison of this sequence as to the recently published genome sequence

CC of C. pneumoniae reveals that the sequence actually contains at least two
CC open reading frames, a first one in the 5' portion and a second one in
CC the 3' portion of the sequence. Despite the presence of the stop codon at
CC the end of this sequence, C. pneumoniae does make a 76 kDa product. It
CC appears possible that C. pneumoniae is able to read through this stop
CC codon and produce a full-length product terminated by the stop codon at
CC the end of the second open reading frame. There is at least one in-frame
CC Arg upstream of the start codon. This suggests that the first open
CC reading frame may form part of one or more larger open reading frames.
CC The lorf2 protein or DNA can be used as a vaccine for humans to treat or
CC prevent disease caused by Chlamydia infection. The sequences or an
CC antibody to lorf2 can be used to diagnose a Chlamydial infection.
CC (Updated on 15-SEP-2003 to standardise OS field)

XX
SQ Sequence 1550 BP; 472 A; 268 C; 317 G; 493 T; 0 U; 0 Other;
Query Match 11.7%; Score 170; DB 3; Length 1550;
Best Local Similarity 100.0%; Pred. No. 4.8e-32;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATTAATAGTCAGAGCTTTTATTTT 60
DB 1381 ATAAATCTTTAAACAGGCTCGCATTAATTAATAGTCAGAGCTTTTATTTT 1440
QY 61 ATAAATCTTTAAACAGGCTCGCATTAATTAATAGTCAGAGCTTTTATTTT 120
DB 1441 ATAAATCTTTAAACAGGCTCGCATTAATTAATAGTCAGAGCTTTTATTTT 1500
QY 121 AGTCTCTATAGAGCAACAGAGCGACACCTCCCGCAGATCTTTCTGCTC 170
DB 1501 AGTCTCTATAGAGCAACAGAGCGACACCTCCCGCAGATCTTTCTGCTC 1550

RESULT 11
AAV16207
ID AAV16207 standard; DNA; 150 BP.
XX
AC AAV16207;
XX

XX 17-OCT-2003 (revised)
XX 28-MAY-1998 (first entry)
XX
XX Part of the gene encoding the 76 kDa cysteine rich OMP.
XX
XX Cysteine rich outer-membrane protein; OMP; 60 kDa OMP; 76 kDa OMP;
XX hybridisation; amplification; assay; detection; ds.
XX
XX Chlamydia pneumoniae.
XX
XX WO9746709-A2.
XX
XX PD 11-DEC-1997.
XX
XX PF 04-JUN-1997; 97WO-US009673.
XX
XX PR 06-JUN-1996; 96US-00659473.
XX
XX PA (ABBO) ABBOTT LAB.
XX
XX PI Cerney MB;
XX
XX DR WPI; 1998-042215/04.
XX
XX FT Chlamydia pneumoniae derived oligonucleotides - used as primers or probes
XX for specific and sensitive detection.
XX
XX PS Disclosure; Page 24; 28pp; English.
XX
XX The present sequence represents part of the gene encoding the 76 kDa
XX cysteine rich outer-membrane protein (OMP) of Chlamydia pneumoniae. Novel
XX oligonucleotides AAV16195-206 can function either as PCR primers or
XX probes. They are used to detect C. pneumoniae in a sample.
XX
XX Oligonucleotides V161202-04 are specific for the present sequence. The

The invention relates to a novel immunogenic composition comprising a protein or nucleic acid, and an adjuvant, where the protein or nucleic acid comprises any of 131 fully defined amino acid or nucleotide sequences given in the specification, or has 50% or greater sequence identity to it, or their fragments. The protein and/or nucleic acid of the immunogenic composition is useful in the manufacture of a medicament for the treatment or prevention of infection due to Chlamydia trachomatis. The infection is treated or prevented by the medicament eliciting an immune response which is specific to a C. trachomatis elementary body, or for neutralising C. trachomatis elementary bodies, hence the immunogenic composition can be used in creating a vaccine. The immunogenic compositions can also be used for the diagnosis of C. trachomatis infection. The nucleic acids of the immunogenic compositions can be used to treat disorders by gene therapy. The immunogenic compositions have antibacterial activity. This polynucleotide sequence represents one of the 131 DNA sequences encoding the C. trachomatis

Query Match	3.7%	Score 54.4	DB 6	Length 1941
Best Local Similarity	45.9%	Pred. No. 0.0038		
Matches 402	Conservative 0	Mismatches 441	Indels 33	Gaps 5
Qy	606	AAAGTTGGCGCAAAATACAGAAATTAGCGAAATATGCTTCGGATACCAAGCGATTCCTG	665	
Db	467	AGGTATTGATAAGCTCATACAGTTAGTTAACTTGAAAAACAGATCAGACTTTAAAGG	526	
Qy	666	ACTCTTAGGTAAACTGACTTCTTCCTTCGACCTCTTTACAGGCTGCTCTCTCCAAATCTCTAG	725	
Db	527	AAACTTTTAAACACACAGACTCTCGAGATCAGATTCCAGCGATTAAATAGTCAGTTAGAGA	586	
Qy	726	CAACAATNACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCAG	785	
Db	587	TCACAACAAAATCTTGCGAGATCAAAATTTACAAGATCTGGAAGGACAAAACATAAGTTATG	646	
Qy	786	GGAACACGCCCTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAGA	845	
Db	647	AAGCTGTTCTCACTTAAACGAGGAGGTTATCAAGCTTCTTCTGAAGCGGAATTAAGT	706	
Qy	846	TAGAGAAGAGTGGAAATCGGANTTAGGATGCATATTTTGCAGGACAGAAACGCTAGTGGAG	905	
Db	707	TAGACAAAGCTTTGCAGTCTATTCTGGATGCTGGGGATCAAGCCAGCTGCAGTCTCTTC	766	
Qy	906	CTGTAGAAAATGCTAAATCTTAATACAGTATATANGCAACATAGATTCAGCTAAACGACAA	965	
Db	767	AAGCACAGCAAAATTAATAGCCAGATTAATTCGCAGCCACAGAAATTAATTGATGCTG	826	
Qy	966	TCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCCAATC	1025	
Db	827	CTGAAAACGAAGGTAAACGAGTTAAAAACAGAGCATACAGGCTTAACGGACTCGCTTTAG	886	
Qy	1026	TTCGAAGACGGGAACAAATGTTAATAACAGCTGTAGAAAGATCTTTAAAAATATCAAACTG	1085	
Db	887	TGAAAAAAGCTGAGGAGCAGATTAGTCAAGCACAAAAAAGATATTCAAGAGATCAAACT	945	
Qy	1086	CAGATGGTCTGTGATGTTCCAAATCCAGGAACATA--CAGTTGGAGGCTCCAGAACAAAGG	1143	
Db	946	--AGTGGTTCGGAATTCCTATCGTTGGTCCGATGGGTACGTGCTTCGCGAGGAAGTG	1003	
Qy	1144	AAGTAGTATTCGTAGTATTC-----GTGTTTCCATGCTGTTTAGATCATG	1187	
Db	1004	CGGTAGGAGCGTTGAAATCCTTAACAAATTCAGGAAGNAITTCCTTGTGCTGATGATG	1063	
Qy	1188	CTGAAAATGAGACCGCTTCCATTTTGATGTCTGGGTTTCGTCAGATGATTCACATGTTCA	1247	

proteins with immunogenic properties of the invention.									
Sequence 1944 BP; 612 A; 391 C; 445 G; 496 T; 0 U; 0 Other;									
Query Match 3.7%; Score 54.4; DB 9; Length 1944;									
Best Local Similarity 45.9%; Pred. No. 0.0039;									
Matches 402; Conservative 0; Mismatches 441; Indels 33; Gaps 5;									
QY	606	AAGTTGGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAACGCAATCTTG	665						
DB	467	AGGTAATGATAAGCTCAATCAGTTAGTTAACTTGAACACAGATCAGACTTTAAAGG	526						
QY	666	ACTCTTTAGTAACTGCTCTTCCTCGACCTCTTACAGGCTGCTCTTCCAACTCTAG	725						
DB	527	AAACITTTAACACACAGACTCTGCAGATCAGATTCACGATTAATAGTCAGTTAGAGA	586						
QY	726	CAACAAATACCAACAGCTGAGCTTCTTAAAGAGATCAAGATAACCCAGTAGTCCCG	785						
DB	587	TCACAAAATTTCTGCAGATCAAAATATCAAAAGATCTGGAAGGACAAACATTAAGT	646						
QY	786	GGAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACAGA	845						
DB	647	AAGCTGTTCTACTAACGACGAGAGGTTATCAAAAGCTTCTTCTGAAGCGGGAATTAAGT	706						
QY	846	TAGAGAAAGATGGAATGCGATTAGGATGCTATTTTCAGGACAGACGCTAGTGGAG	905						
DB	707	TAGGACAGCTTTGCAGTCTATTGTGGATGCTGGGGATCAAGCCAGGCTGCAGTCTTC	766						
QY	906	CTGTAGAAATCTAAATCTTAATAACAGTATTAAGCAACATAGATTTCAGCTAAAGCAGAA	965						
DB	767	AAGCAGACAAATTAATAGCCAGATTAATTCGAGCCACAGGAATAATTAATTGATGCTG	826						
QY	966	TCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAATTC	1025						
DB	827	CTGAACGAGGATAAACGAGTTAAACAAAGAGCATACAGGCTTAACGGACTCGCTTTAG	886						
QY	1026	TTCAAGACGGGACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCAAACTTG	1085						
DB	887	TGAAAAAGCTGAGGAGCAGATTAGTCAAGCACAATAAGATATTCAGAGATCAAACTT-	945						
QY	1086	CAGATGTTCTGATGTTCCAAATCCAGGAACATA--CAGTTGGAGGCTCCAGCAACAGG	1143						
DB	946	--AGTGGTTCGGATATCTCTATCGTTGGTCCGAGTGGTTCAGCTGTTCCGAGGAAGTG	1003						
QY	1144	AAGTAGTATTTGGTAGTATTC-----GTGTTTCCATGCTGTTAGATGATG	1187						
DB	1004	CGGTAGGAGCGTTGAAATCCTCTPAACAAATTCAGGAAGAAATTCCTTGTGCTTGATGAG	1063						
QY	1188	CTGAAATGAGACCGCTTCCATTTTGTATGTTCTGGGTTTCGTACAGATGATTCACATGTTCA	1247						
DB	1064	TAGACAAATGAATGGCAGCGAATTCGAATGCAAGTTTTCGATCTATGATCGAACAAATTA	1123						
QY	1248	ATACGGAATATCTGATTTCTCAAGCTGCCAACAGGAGCTTCGACGACCAAGCTTAGAGCAG	1307						
DB	1124	ATGTAACATCTCTGCAACAGCTTAAGAGCTACAGCTATGGAGGCTCAGCTGA---CTG	1180						
QY	1308	CGAAGCCGCTGGAGATGACAGTGTCTGCTCAGCGCTGGCAGATGCTCAGAAAGCTTTAG	1367						
DB	1181	CGATGTCAGATCAACTGTTTGGTGGGATGGGAGTCCCGAGCTCCCGAGGAAATCAAGCAATCA	1240						
QY	1368	AAGGGCTCTAGTAAAGCTGGGCAACAC-----AGGCACTCAATCAATGCTTTAG	1418						
DB	1241	AAGATGCTCTTGGCCAAAGCTTTGAAACAAACCATCAACAGATGGTTTAGCTACAGTATGG	1300						
QY	1419	GACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAG	1454						
DB	1301	GACAGTGGCTTTTGCAGCTGCCAAGGTTGGAGGAG	1336						

Search completed: March 23, 2004, 20:03:04

Job time : 563.656 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:22:08 ; Search time 100.993 Seconds
(without alignments)
8000.664 Million cell updates/sec

Title: US-10-608-559-5
Perfect score: 1456
Sequence: 1 ataaatctttaaaacagg.....ctgttgagcaggaagta 1456

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1444	99.2	1230025	4	US-09-198-452A-1
2	170	11.7	1550	4	US-09-427-501-1
3	170	11.7	1550	4	US-09-905-119A-1
4	148.4	10.2	150	3	US-08-659-473-8
5	54.4	3.7	7218	1	US-08-232-463-14
6	51.6	3.5	2205	4	US-09-328-352-286
7	51.4	3.5	1537	4	US-09-556-877-57
8	51.4	3.5	1537	4	US-09-620-412C-57
9	51.4	3.5	1537	4	US-09-410-568-57
10	51.4	3.5	1537	4	US-09-598-419-57
11	49	3.4	1990	4	US-08-714-741-35
12	49	3.4	2085	1	US-08-072-070-1
13	49	3.4	2085	1	US-08-465-746-1
14	49	3.4	2085	1	US-08-214-164-1
15	49	3.4	2085	1	US-08-469-434-1
16	49	3.4	2085	1	US-08-214-222-1
17	49	3.4	2085	2	US-08-467-852A-1
18	49	3.4	2085	2	US-08-468-719-1
19	49	3.4	2085	2	US-08-246-636-1
20	49	3.4	2085	2	US-08-247-491A-1
21	49	3.4	2085	2	US-08-319-795-1
22	49	3.4	2085	2	US-08-468-985-1
23	49	3.4	2086	3	US-08-312-949-1
24	49	3.4	2086	3	US-08-446-201-2
25	47.4	3.3	956	4	US-08-714-741-36
26	43.6	3.0	6744	1	US-08-119-125A-2
27	39.8	2.7	1295	4	US-09-286-981B-20

28	39.8	2.7	7953	4	US-08-956-171E-152
29	39.6	2.7	2208	4	US-09-134-001C-603
30	39.2	2.7	1230025	4	US-09-198-452A-1
C 31	39	2.7	2550	6	5258287-23
C 32	38.8	2.7	1866	4	US-09-601-198-153
33	38.6	2.7	8961	4	US-10-204-708-80
34	38.4	2.6	1196	4	US-08-956-171E-71
35	38.4	2.6	2598	4	US-09-612-204B-23
C 36	38.4	2.6	1664976	4	US-08-916-421B-1
37	38.2	2.6	1254	4	US-09-540-236-547
38	38	2.6	471	4	US-09-134-001C-2415
39	38	2.6	640	3	US-08-961-083-33
40	38	2.6	640	4	US-09-536-784-33
41	38	2.6	6314	4	US-09-620-312D-98
42	38	2.6	6583	4	US-10-204-708-26
43	38	2.6	30549	4	US-09-134-001C-322
44	37.8	2.6	1470	4	US-09-134-001C-2404
C 45	37.8	2.6	3337	1	US-08-072-610-1

ALIGNMENTS

RESULT 1

US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (150001)..(165000)

Sequence 152, App
Sequence 603, App
Sequence 1, Appli
Patent No. 5258287
Sequence 153, App
Sequence 80, Appl
Sequence 71, Appl
Sequence 23, Appl
Sequence 1, Appli
Sequence 547, App
Sequence 2415, App
Sequence 33, Appl
Sequence 33, Appl
Sequence 98, Appl
Sequence 26, Appl
Sequence 322, App
Sequence 2404, App
Sequence 1, Appli


```
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature

Query Match
Best Local Similarity 99.2%; Score 1444; DB 4; Length 1230025;
Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATAAATCTTTAAAAACAGGCTCGCATTAATATTATCTAGTGGAGGCTTTTATTTTTT 60
DB 828698 ATAAATCTTTAAAAACAGGCTCGCATTAATATTATCTAGTGGAGGCTTTTATTTTTT 828639

QY 61 ATAAATCTTTAAAAACAGGCTCGCATTAATATTATCTAGTGGAGGCTTTTATTTTTT 120
DB 828638 ATAAATCTTTAAAAACAGGCTCGCATTAATATTATCTAGTGGAGGCTTTTATTTTTT 828579

QY 121 AGTCTTATAGCAAAACAGAACGCAACCTCCGCGAGATCTTCTCTCAAGATTTGA 180
DB 828578 AGTCTTATAGCAAAACAGAACGCAACCTCCGCGAGATCTTCTCTCAAGATTTGA 828519

QY 181 GCGAGTGCAGCAAAATAGAGTGGGAAGCTCRAAGAAATAGCAGGTGCGGAAGCTAAGCC 240
DB 828518 GCGAGTGCAGCAAAATAGAGTGGGAAGCTCRAAGAAATAGCAGGTGCGGAAGCTAAGCC 828459

QY 241 TAAAGAACTTAAGACCGATCTGTAGAGCGATGAGGATCTTGGTTCTCGAGTGAATGC 300
DB 828458 TAAAGAACTTAAGACCGATCTGTAGAGCGATGAGGATCTTGGTTCTCGAGTGAATGC 828399

QY 301 TCTCATGAGTCTGCAGATTAAGTGGGTATGCTTCTAGTAAACAGTCTGTTCTACTAG 360
DB 828398 TCTCATGAGTCTGCAGATTAAGTGGGTATGCTTCTAGTAAACAGTCTGTTCTACTAG 828339

QY 361 CAGATCTGCAGAGCTGACTCAACGACAGCGACCGACCTAGGCTCTCTCCACCAGTT 420
DB 828338 CAGATCTGCAGAGCTGACTCAACGACAGCGACCGACCTAGGCTCTCTCCACCAGTT 828279

QY 421 TGATGATTATTAAGACTCAAGCGCAAAACAGCTTACGATPACTATCTTTACCTCAACATCACT 480
DB 828278 TGATGATTATTAAGACTCAAGCGCAAAACAGCTTACGATPACTATCTTTACCTCAACATCACT 828219

QY 481 AGTGCATACAGGCTCTTGTGTAGAGCTCCAGGATGCTGTCACTAATATAAGATAC 540
DB 828218 AGTGCATACAGGCTCTTGTGTAGAGCTCCAGGATGCTGTCACTAATATAAGATAC 828159

QY 541 AGCGGCTPACTGATGAGGAACCGCAATCGTGTGCGAGTGGGAACTAAGAAATGCCGATGC 600
DB 828158 AGCGGCTPACTGATGAGGAACCGCAATCGTGTGCGAGTGGGAACTAAGAAATGCCGATGC 828099

QY 601 AGTTAAAGTTGGCGCGCAAAATTAAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGAT 660
DB 828098 AGTTAAAGTTGGCGCGCAAAATTAAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGAT 828039

QY 661 TCTTGACTCTTTAGTAAATGACTTCTCTGACCTCTTACGCTCTCTTCTTCAATC 720
DB 828038 TCTTGACTCTTTAGTAAATGACTTCTCTGACCTCTTACGCTCTCTTCTTCAATC 827979

QY 721 TGTAGCAAAACAAATAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGT 780
DB 827978 TGTAGCAAAACAAATAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGT 827919

QY 781 CCCAGGAAAACCGCTCGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
DB 827918 CCCAGGAAAACCGCTCGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 827859

QY 841 ACAGATAGAAAGATGGAATGCGATAGGGATGCAATTTTCCAGGACAGAACGCTAG 900
DB 827858 ACAGATAGAAAGATGGAATGCGATAGGGATGCAATTTTCCAGGACAGAACGCTAG 827799

QY 901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAAGCTAAGC 960
DB 827798 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAAGCTAAGC 827739

QY 961 AGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
DB 827738 AGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 827679

QY 1021 AATCTTCAAGAACGGAACAATGTAATACAGGCTGAGAAAGATCTTAAATATATCAA 1080
DB 827678 AATCTTCAAGAACGGAACAATGTAATACAGGCTGAGAAAGATCTTAAATATATCAA 827619

QY 1081 ACCTGCAGATGGTCTCTGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACA 1140
DB 827618 ACCTGCAGATGGTCTCTGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACA 827559

QY 1141 AGGAATAGTATTGGTAGTATTCTGTTTCCATGCTTTAGATGATGCTGAAAATGAGAC 1200
DB 827558 AGGAATAGTATTGGTAGTATTCTGTTTCCATGCTTTAGATGATGCTGAAAATGAGAC 827499

QY 1201 CGTTTCCATTTTGCATGCTGGGTTTTCGTAGATGATTCATGTTCAATACGAAAATCC 1260
DB 827498 CGTTTCCATTTTGCATGCTGGGTTTTCGTAGATGATTCATGTTCAATACGAAAATCC 827439

QY 1261 TGATTTCAAGCTGCCCAACAGGAGCTCGCAGCACCAAGCTTAGAGCAGCGAAAGCGCTGG 1320
DB 827438 TGATTTCAAGCTGCCCAACAGGAGCTCGCAGCACCAAGCTTAGAGCAGCGAAAGCGCTGG 827379

QY 1321 AGATGACGTGCTGTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTTAGG 1380
DB 827378 AGATGACGTGCTGTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTTAGG 827319

QY 1381 TAAAGCTGGCAACAAACAGGCTACTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTG- 1439
DB 827318 TAAAGCTGGCAACAAACAGGCTACTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTG- 827259

QY 1440 TTGTGAGCGCAGGAGT 1455
DB 827258 TTGTGAGCGCAGGAGT 827243

RESULT 2
US-09-427-501-1
; Sequence 1, Application US/09427501
; Patent No. 6403101
; GENERAL INFORMATION:
; APPLICANT: Connaught Laboratories Limited, Andrew
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 19721-010
; CURRENT APPLICATION NUMBER: US/09/427,501
; EARLIER FILING DATE: 1999-10-26
; EARLIER FILING DATE: 1998-10-28
; EARLIER FILING DATE: 1999-09-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1550
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1366)
US-09-427-501-1

Query Match
Best Local Similarity 11.7%; Score 170; DB 4; Length 1550;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAAAACAGGCTCGCATTAATATTATCTAGTGGAGGCTTTTATTTTTT 60
DB 1381 ATAAATCTTTAAAAACAGGCTCGCATTAATATTATCTAGTGGAGGCTTTTATTTTTT 1440
```



```
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match          3.7%; Score 54.4; DB 1; Length 7218;
Best Local Similarity 9.3%; Pred. No. 0.00017;
Matches 43; Conservative 218; Mismatches 199; Indels 0; Gaps 0;

Qy 731 AATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGTCCCGAGGAAA 790
    |||
Db 1481 AATTACCTATCTATGCAAGTAGTTTAAAGAGATAGAGAAATTTGCTACRRRRRRRRRR 1422

Qy 791 ACGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACACAGATAGAG 850
    |||
Db 1421 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1362

Qy 851 AAAGATGAAATGCGATTAGGATGTCATATTTCGAGCAGACAGACGCTAGTGGAGCTGTA 910
    |||
Db 1361 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1302

Qy 911 GAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGAGCTTAAAGCAGCAATCGCT 970
    |||
Db 1301 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1242

Qy 971 ACTGCTAAGACAAATAGCTGAGCTCAGAAAAGTTCCCGACTCTCCAATCTTCAA 1030
    |||
Db 1241 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1182

Qy 1031 GAAGCGCAACAAATGGTAAATACAGCTGAGAAAGATCTTAAATAATCAAAACCTGCAGAT 1090
    |||
Db 1181 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1122

Qy 1091 GGTCTGATGTTCCAAATCCAGGAACACTAGCTGGAGGCTCCAGCAACAGGAAGTAGT 1150
    |||
Db 1121 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1062

Qy 1151 ATTGCTAGTATTCTGGTTTCCATGCTGTTAGATGATGCTG 1190
    |||
Db 1061 AAGCTCCCTCGACCTGCAGCAAGCTCGGAATTAATCTG 1022

RESULT 6
US-09-328-352-286
; Sequence 286, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 286
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-286

Query Match          3.5%; Score 51.6; DB 4; Length 2205;
```

```
Best Local Similarity 54.8%; Pred. No. 0.00052;
Matches 102; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 791 ACGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACACAGATAGAG 850
    |||
Db 1426 ACAATTTCTATTTCACCCAGCTTTAAACGATAATAAATCAATGCTTCGGTATCAGCTATAGAT 1485

Qy 851 AAAGATGGAATGCGATTAGGATGTCATATTTCGAGCAGACAGACGCTAGTGGAGCTGTA 910
    |||
Db 1486 AATGCTGGAATAAGTCTGAAGTTGTTGATTTAGGTACAAAGAGATACAACACCA 1545

Qy 911 GAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGAGCTAAGCAGCAATCGCT 970
    |||
Db 1546 GCAAAACCTATATTAATAGCTAGATGATGATGCTGCTTTAAAGGAGCTATAACG 1605

Qy 971 ACTGCT 976
    |||
Db 1606 GCTGGT 1611

RESULT 7
US-09-556-877-57
; Sequence 57, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 57
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
; US-09-556-877-57

Query Match          3.5%; Score 51.4; DB 4; Length 1537;
Best Local Similarity 53.5%; Pred. No. 0.00049;
Matches 159; Conservative 0; Mismatches 126; Indels 12; Gaps 2;

Qy 1167 TTTCATGCTGTTAGATGATGCTGAAATGAGACCGCTTCCATTTTGATGCTGGGTTTC 1226
    |||
Db 186 TTTCCTTTGCTGTGATGATGATGACCAATGAAATGGCAGCGATGCAATGCAAGGTTTC 245

Qy 1227 GTCAGATGATTCATGTTCAATACGGAAATCTGATTTCAAGCTGCCCAACAGGAGC 1286
    |||
Db 246 GATCTATGATCGAACAAATTTAATGTAACCAATCTCGAACAGCTTAAAGAGGCTACAAGCTA 305

Qy 1287 TCGCAGCACAGCTAGAGCAGCAAGCAAGCGCTGAGATGACAGTGTCTGACGGCTGG 1346
    |||
Db 306 TGGAGGCTCAGCTGA---CTGCGATGTCAGATCAACTGTTGTTGGGATGGCGAGCTCC 362

Qy 1347 CAGATGCTCAGAAAGCTTTTGAAGCGCTCTAGTAAAGCTGGGCAACAAC- 1397
    |||
Db 363 CAGCGAAATACAGCAATCAAGATGCTCTTGGCAAGCTTTTGAACAACCATCAGCAG 422

Qy 1398 AGGCATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGTGTGTGAGCGCAGGAG 1454
    |||
Db 423 ATGCTTTAGCTACAGCTATGCGACAGTGGCTTTTTCAGCTGCCCAAGGTTGGAGGAG 479

RESULT 8
US-09-620-412C-57
; Sequence 57, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
```

Db	246	GATCTATGTCGAACAAATTAAATGTAACAACTCTGCAACAGCTAAAGAGCTACAAGCTA	305
Qy	1287	TCCGAGCACAAGCTTAGAGCAGCGAAGCCGCTGGAGATGACAGTGTCTGTGACGCGCTGG	1346
Db	306	TGGAGGCTCAGCTGCA---CTGCCATGTCAGATCAACTGTTGGTGGGATGGCGAGCTCC	362
Qy	1347	CAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGTAAAGCTGGGGCAACAC--	1397
Db	363	CAGCGAATACAGCAATCAAGATGCTCTTGGCAAGCTTTGAACAACCATCAGCAG	422
Qy	1398	AGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGTGTGTGAGCGCAGGAG	1454
Db	423	ATGGTTAGCTACAGCTATGCGACAAGTGGCTTTTGAGCTGCCAAGGTTGGAGGAG	479

```

RESULT 10
US-09-598-419-57
; Sequence 57, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Schellier, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 57
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-57

```

QY	1167	TTTCCATGCTGTTAGATGATGCTCAAAAATAGACCGGCTTCATTTTGATGCTGCTGGTTTC	1226
DB	186	TTTCCCTTTGCTTGATGATGTAGCAATGAAATGGCAGCGATTGCAATGCAGGTTTTC	245
QY	1227	GTCAGATGATTCAATGTTCAATACGGAAAATCCTGATTCTCAAGCTGCCCAACAGGAGC	1286
DB	246	GATCTATGATCGAACAAATTTAATGTAACCAATCTGCACACAGCTAAAGAGGTACAAGCTA	305
QY	1287	TCGCAGACACAGCTTAGACGACGCGGCTGGAGATGACAGTGCTGCTCAGCGCTGG	1346
DB	306	TGGAGGCTCAGCTGA--CTCGATGTCAGATCAACTGTTGGTGGGATGCGGAGCTCC	362
QY	1347	CAGATGCTCAGAAAGCTTTAAGCGGCTTAGGTAAAGCTGGCGCAACAC	1397
DB	363	CAGCCGAANTACAAGCAATCAAGATGCTCTGTGCGCAAGCTTTGAAACACCATCAGCAG	422
QY	1398	AGGCGTACTCAATGCTTTTAGCAGAGATCGCTTCTGCTGCTGTTGTCGAGCGCAGGAG	1454
DB	423	ATGCTTAGCTTACAGCTATGGAACAGTGGCTTTTTCAGCTGCGCAAGCTTGGAGAG	479

Patent No. 6500613
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: McDaniel, Harry S.
APPLICANT: Swiatlo, Edwin
APPLICANT: Yotter, Janet
APPLICANT: Crain, Marilyn J.
APPLICANT: Hollingshead, Susan
APPLICANT: Tart, Rebecca

;; APPLICANT: Brooks-Walter, Alexis
;; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
;; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
;; TITLE OF INVENTION: PORTIONS AND PRODUCTS
;; NUMBER OF SEQUENCES: 47
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Curtis, Morris & Safford, P.C.
;; STREET: 530 Fifth Avenue
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.
;; ZIP: 10036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/714,741
;; FILING DATE: 16-SEP-1996
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Frommer Esq., William S.
;; REGISTRATION NUMBER: 25,506
;; REFERENCE/DOCKET NUMBER: 454312-2460
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 840-3333
;; TELEFAX: (212) 840-0712
;; INFORMATION FOR SEQ ID NO: 35:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1990 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-714-741-35

Query Match 3.4%; Score 49; DB 4; Length 1990;
Best Local Similarity 49.5%; Pred. No. 0.0025;
Matches 155; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

Qy 845 ATAGAGAAAGATGGAATGCGATAGGATGCATATTTTCAGGACAGACGCTAGTGG 904
Db |||||
472 ACAGACAAAGCGCGAAAGACGACAGATAGATAGATAGCTTAAGAAACGGAA 531
Qy |||||
905 GCTGTAGAAATGCTTAATCTAATTAACAGTATAGCAACATAGATTCTAGCTAAAGCAGCA 964
Db |||||
532 GAAGAGGCAAACTTAATTTTAATCTGTCGACATGGTAGTTCTCTGAGCCAGACG 591
Qy |||||
965 ATCGCTACTGTAAAGACAAATAGCTGAAGCTCAGAAAAAGTTCCCGA---CTCTCCA 1021
Db |||||
592 TTGCTGAGCTAGGAAATTAACAGAAAGCTTAACAAAAGCAACCAAGCACTTACTTAA 651
Qy |||||
1022 ATTTCTAAGAGCGGACAAATAGTATACAGCTGAGAAGATCTTTAAAAATATCAAA 1081
Db |||||
652 AAACCTAGAAGAGCTTAAGCAAAATTAAGAGGCTGAGAAAAAGCTTACTGAAGCCAAA 711
Qy |||||
1082 CTGCAAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGAGGGTCCCAAGCAAA 1141
Db |||||
712 CAAAAGTGGATGCTGAGAAAGTCGCTCTCAAGCTAAATCGCTGAAATTCGGAATCA 771
Qy |||||
1142 GGAAGTATGTTG 1154
Db |||||
772 GTTCATAGACTAG 784

RESULT 12
US-08-072-070-1
; Sequence 1, Application US/08072070
; Patent No. 5476929
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L

;; APPLICANT: McDaniel, Larry S
;; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
;; TITLE OF INVENTION: PROTEIN
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Shoemaker and Mattare, Ltd
;; STREET: Suite 1203, 2001 Jefferson Davis Highway
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202-0286
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/072,070
;; FILING DATE: 19930603
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/835,698
;; FILING DATE: 12-FEB-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/656,773
;; FILING DATE: 15-FEB-1991
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 415-0810
;; TELEFAX: (703) 521-0378
;; TELEX: LUKPAT WASHINGTON
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2085 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: Gouble
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Streptococcus pneumoniae
;; STRAIN: Rx1
;; IMMEDIATE SOURCE:
;; CLONE: JY2008
;; FEATURE:
;; NAME/KEY: intron
;; LOCATION: 1..2085
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: join(127..1983, 1987..1992, 1996..2007, 2011
;; LOCATION: ..2025, 2029..2031, 2035..2085)
;; US-08-072-070-1

Query Match 3.4%; Score 49; DB 1; Length 2085;
Best Local Similarity 49.5%; Pred. No. 0.0025;
Matches 155; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

Qy 845 ATAGAGAAAGATGGAATGCGATAGGATGCATATTTTCAGGACAGACGCTAGTGG 904
Db |||||
472 ACAGACAAAGCGCGAAAGACGACAGATAGATAGATAGCTTAAGAAACGGAA 531
Qy |||||
905 GCTGTAGAAATGCTTAATCTAATTAACAGTATAGCAACATAGATTCTAGCTAAAGCAGCA 964
Db |||||
532 GAAGAGGCAAACTTAATTTTAATCTGTTTCGAGCAATGGTAGTTCTCTGAGCCAGCAG 591
Qy |||||
965 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGA---CTCTCCA 1021
Db |||||
592 TTGCTGAGACTAAGAAAAAATCAGAGAGAGCTAAACAAAAAGCACCAGAACTTACTAAA 651
Qy |||||
1022 ATTTCTAAGAGCGGACAAATGGTAAATACAGCTGAGAAGATCTTTAAAAATATCAAA 1081
Db |||||
652 AAACCTAGAAGAGCTTAAGCAAAATTAAGAGGCTGAGAAAAAGCTTACTGAAGCCAAA 711

QY	845	ATAGAGAAAAGATGCAAAATCGAATAGGGATGCAATATTTTCAGACACAGAAACGCTAGTGGAA	904
Db	472	ACAGACAAAGCCGCAAAAAGCGCAGCAGATAAAGATGATAGATGAAGCTAAGAAAACGGCGAA	531
QY	905	GCTGTGAAAATGCTAAATCTTATAACAGTATTAAGCAACATAGATTTCAGCTTAAGCAGCA	964

Qy	965	ATCGCTACTGCTTAAGACACAAAATAGCTGTGAAGCTCAGAAAAAGATCTCCUGA---CTCTCCA	108
Db	592	TTGCGCTGAGACTAAGAAAAAATCAGAGAGAAGCTAAACAAAAGACCCAGAGAACTTACTATAA	651
Qy	1022	ATTCTTCAAGAAGCGGAACAAAATGGTAAATACAGGCTGAGAGAAGATCTTTAAAAAATATCAAA	1088
Db	652	AAACTAGAGAAGCTAAAGCATAAATTAAGAAGAGAGGCTGAGAAAAAGGCTACTTGAAGCCAAA	711
Qy	1082	CCTTGCAGATGGTCTTGATGTTCCAAATCCAGGAAGAACTACAGTTGGAGGCTCCCAAGCAACAA	1147
Db	712	CAAAAAGTGATGCTGAAGAAAGTCGTCTCTCAAGCTAAATCATCGTGAATTCGAAATCAAA	771

Qy 1142 GGAAGTAGTATTG 1154
 |||||
 Db 772 GTTCATAGACTAG 784

RESULT 14

US-08-214-164-1
; Sequence 1, Application US/08214164
; Patent No. 5728387
; GENERAL INFORMATION:
; APPLICANT: BRILES, DAVID E.
; APPLICANT: YOTNER, JANET L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN

CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd
STREET: Suite 1203, 2001 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.

```
; ZIP: 22202-0286  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/214,164
: FILING DATE: 17-MAR-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/656,773
: FILING DATE: 15-FEB-1997
:

```

FILING DATE: 15-FEB-1991
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Berkstresser, Jerry W.
 REGISTRATION NUMBER: 22,651
 REFERENCE/DOCKET NUMBER: 6102-137
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 415-0810
TELEFAX: (703) 521-0813
TELEX: LUKPAT WASHINGTON
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2085 base pairs

```

NAME: ...
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```

```
;
; FEATURE: intron
; NAME/KEY:
; LOCATION: 1..1983
;
```

```
;
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 127..1983
; US-08-214-164-1

Query Match 3.4%; Score 49; DB 1; Length 2085;
Best Local Similarity 49.5%; Pred. No. 0.0025;
Matches 155; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

Qy 845 ATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTTCAGGACAGAACGCTAGTGGG 904
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
472 ACAGACAAAGCCGCAAGAGCGCAGATAGATGATGATGATGATGATGATGATGATGATGAT 531
Qy 905 GCTGTAGAAAATGCTAAATCTAAATCTAAATCTAAATCTAAATCTAAATCTAAATCTAA 964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
532 GAAGAGGCAAAACTAAATTTTAACTAAATTTTAACTAAATTTTAACTAAATTTTAACTAA 591
Qy 965 ATGCTACTGCTAAGACACAATAGCTGAAGCTCAGAAAAAGTTCCCGCA---CTCTCCA 1021
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
592 TTGCTGAGACTAAGAAAAATCAGAGAGCTTAACAAAAAGCACCAGAACTTACTAAA 651
Qy 1022 ATTCTTCAAGAGCGGCAACAAATGTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1081
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
652 AAACCTAGAGAGGCTAAAGCAAAATTTAGAGAGGCTGAGAAAAAGCTACTGAAAGCCAAA 711
Qy 1082 CCTCAGATGTTCTGTATGTTCCAAATCCAGGAAGTACAGTTGAGGCTCCAGCAACAA 1141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
712 CAAAAGTGGATGCTGAAGAAAGTCGCTCCTCAAGCTAAATTCGCTGAATTTGGAANAATCAA 771
Qy 1142 GGAAGTAGTATTG 1154
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
772 GTTCATAGACTAG 784

RESULT 15
US-08-469-434-1
; Sequence 1, Application US/08469434
; Patent No. 5753463
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,434
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,065
; FILING DATE: 03 JUNE 1993
; APPLICATION NUMBER: US/07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
```

```
;
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2085 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: Rx1
; IMMEDIATE SOURCE:
; CLONE: JY2008
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1..2085
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(127..1983, 1987..1992, 1996..2007, 2011
; LOCATION: ..2025, 2029..2031, 2035..2085)
; US-08-469-434-1

Query Match 3.4%; Score 49; DB 1; Length 2085;
Best Local Similarity 49.5%; Pred. No. 0.0025;
Matches 155; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

Qy 845 ATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTTCAGGACAGAACGCTAGTGGG 904
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
472 ACAGACAAAGCCGCAAGAGCGCAGATAGATGATGATGATGATGATGATGATGATGATGAT 531
Qy 905 GCTGTAGAAAATGCTAAATCTAAATCTAAATCTAAATCTAAATCTAAATCTAAATCTAA 964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
532 GAAGAGGCAAAACTAAATTTTAACTAAATTTTAACTAAATTTTAACTAAATTTTAACTAA 591
Qy 965 ATGCTACTGCTAAGACACAATAGCTGAAGCTCAGAAAAAGTTCCCGCA---CTCTCCA 1021
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
592 TTGCTGAGACTAAGAAAAATCAGAGAGCTTAACAAAAAGCACCAGAACTTACTAAA 651
Qy 1022 ATTCTTCAAGAGCGGCAACAAATGTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1081
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
652 AAACCTAGAGAGGCTAAAGCAAAATTTAGAGAGGCTGAGAAAAAGCTACTGAAAGCCAAA 711
Qy 1082 CCTCAGATGTTCTGTATGTTCCAAATCCAGGAAGTACAGTTGAGGCTCCAGCAACAA 1141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
712 CAAAAGTGGATGCTGAAGAAAGTCGCTCCTCAAGCTAAATTCGCTGAATTTGGAANAATCAA 771
Qy 1142 GGAAGTAGTATTG 1154
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
772 GTTCATAGACTAG 784
```

Search completed: March 23, 2004, 13:39:19
Job time : 105.993 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:23:54 ; Search time 502.565 Seconds
(without alignments)
10716.954 Million cell updates/sec

Title: US-10-608-559-5

Perfect score: 1456

Sequence: 1 ataaatctttaaaacagg.....ctgttgagcaggagtagta 1456

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubna/PCNT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubna/PCNT_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1456	100.0	1456	10	US-09-564-479-5	Sequence 5, Appli
2	1456	100.0	2238	10	US-09-564-479-7	Sequence 7, Appli
3	1455	99.9	2156	10	US-09-564-479-4	Sequence 1, Appli
4	1444	99.2	1230025	15	US-10-289-762-1	Sequence 1, Appli
5	1355	93.1	1956	9	US-09-841-132-385	Sequence 385, App
6	1355	93.1	1956	10	US-09-841-260-63	Sequence 63, Appl
7	1355	93.1	1956	15	US-10-007-693-63	Sequence 63, Appl
8	1355	93.1	1956	15	US-10-312-273-14	Sequence 14, Appl
9	1151	79.1	1852	10	US-09-564-479-3	Sequence 3, Appli
10	170	11.7	1550	9	US-09-905-119-3	Sequence 1, Appli
11	54.4	3.7	1941	9	US-09-841-132-412	Sequence 412, App
12	51.4	3.5	1537	9	US-09-841-132-57	Sequence 57, Appl
13	49	3.4	1860	12	US-10-282-122A-37518	Sequence 37518, A
14	48.6	3.3	1171	10	US-09-841-260-13	Sequence 13, Appl
15	48.6	3.3	1171	13	US-10-007-693-13	Sequence 13, Appl

16	48.6	3.3	1834	10	US-09-841-260-38	Sequence 38, Appl
17	48.6	3.3	1834	13	US-10-007-693-38	Sequence 38, Appl
18	48.6	3.3	1983	10	US-09-841-260-28	Sequence 28, Appl
19	48.6	3.3	1983	13	US-10-007-693-28	Sequence 28, Appl
20	47	3.2	9373	12	US-10-221-613-178	Sequence 178, App
21	47	3.2	9373	14	US-10-311-455-1264	Sequence 1264, App
22	44.4	3.0	2499	9	US-09-842-552-92	Sequence 92, Appl
23	44.2	3.0	4997	12	US-10-282-122A-35506	Sequence 35506, A
24	44	3.0	3673778	14	US-10-312-841-2	Sequence 2, Appli
25	43.8	3.0	702	15	US-10-027-632-45880	Sequence 45880, A
26	43.8	3.0	702	15	US-10-027-632-45881	Sequence 45881, A
27	43.5	3.0	6341	14	US-10-311-455-1617	Sequence 1617, App
28	43	3.0	18218	14	US-10-311-455-1922	Sequence 1922, App
29	43	3.0	18585	14	US-10-240-485-161	Sequence 161, App
30	43	3.0	3673778	14	US-10-312-841-1	Sequence 1, Appli
31	42.8	2.9	302	12	US-10-424-599-63444	Sequence 63444, A
32	42.8	2.9	2404	15	US-10-104-047-1740	Sequence 1740, App
33	42.8	2.9	3439	15	US-10-161-927-53	Sequence 53, Appl
34	42.8	2.9	16688	14	US-10-311-455-293	Sequence 293, App
35	42.6	2.9	6126	14	US-10-311-455-1418	Sequence 1418, App
36	42.4	2.9	330	10	US-09-814-353-5359	Sequence 5359, App
37	42.4	2.9	330	10	US-09-814-353-11646	Sequence 11646, A
38	42.4	2.9	5503	16	US-10-257-166-104	Sequence 104, App
39	42.2	2.9	6692	12	US-10-221-714A-131	Sequence 131, App
40	42	2.9	6535	14	US-10-311-455-903	Sequence 909, App
41	41.8	2.9	5278	12	US-10-221-714A-97	Sequence 97, Appl
42	41.8	2.9	5278	14	US-10-311-455-795	Sequence 795, App
43	41.8	2.9	6338	14	US-10-311-455-1099	Sequence 1099, App
44	41.6	2.9	488	10	US-09-918-995-31590	Sequence 31590, A
45	41.6	2.9	576	10	US-09-814-353-5208	Sequence 5208, App

ALIGNMENTS

RESULT 1
US-09-564-479-5
; Sequence 5, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COHEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1456
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1456)
US-09-564-479-5

Query Match	100.0%;	Score 1456;	DB 10;	Length 1456;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1456;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATAAAATCTTTAAAAACAGGCTGCATTAATTTATTAGTAGAGAGCTTTTATTTT	60	
Db	1	ATAAAATCTTTAAAAACAGGCTGCATTAATTTATTAGTAGAGAGCTTTTATTTT	60	
Qy	61	ATAATAAACTAAAGATTTTATTTTATTTTATGTTTATGTTTATGTTTATGTTT	120	

Db 906 TAAAGAACTAAGACCGATTCGTAGAGCGATGGAGCATCTTGGCTTCGAGTGAATGC 965
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGTATTCCTTCTAGTAACAGCTCGCTCTTCTACTAG 360
Db 966 TCTCATGAGTCTGGCAGATAGCTGGTATTCCTTCTAGTAACAGCTCGCTCTTCTACTAG 1025
QY 361 CAGATCTGAGACGTGGAGTCAACGACGACCGCCTACGCTCTCCACCCACGTT 420
Db 1026 CAGATCTGAGACGTGGAGTCAACGACGACCGCCTACGCTCTCCACCCACGTT 1085
QY 421 TCATGATTAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT 480
Db 1086 TCATGATTAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT 1145
QY 481 AGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGCACCTAAATATAAGGATAC 540
Db 1146 AGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGCACCTAAATATAAGGATAC 1205
QY 541 AGCGGCTACTGATGAGGAACCGCAATCGCTGGGAGTGGGAAACTAAGAAATGCCGATGC 600
Db 1206 AGCGGCTACTGATGAGGAACCGCAATCGCTGGGAGTGGGAAACTAAGAAATGCCGATGC 1265
QY 601 AGTTAAAGTTGGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT 660
Db 1266 AGTTAAAGTTGGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT 1325
QY 661 TCTTGACTCTTTAGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCCAAATC 720
Db 1326 TCTTGACTCTTTAGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCCAAATC 1385
QY 721 TGTAGCAAAATATAAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db 1386 TGTAGCAAAATATAAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 1445
QY 781 CCAGGGAACCCCTGCAATTCCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGAC 840
Db 1446 CCAGGGAACCCCTGCAATTCCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGAC 1505
QY 841 ACAGATAGAAAGATGGAAATCGATTAGGATGCATATTTTGCAGGACAGAACGCTAG 900
Db 1506 ACAGATAGAAAGATGGAAATCGATTAGGATGCATATTTTGCAGGACAGAACGCTAG 1565
QY 901 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGC 960
Db 1566 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGC 1625
QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
Db 1626 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1685
QY 1021 AATTTCTAAGAGCGGAACAAATGTTAATAACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
Db 1686 AATTTCTAAGAGCGGAACAAATGTTAATAACAGGCTGAGAAAGATCTTAAAAATATCAA 1745
QY 1081 ACTGCGAGTGGTCTGATGTTCCCAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 1140
Db 1746 ACTGCGAGTGGTCTGATGTTCCCAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 1805
QY 1141 AGGAAGTAGTATTGGTAGTATTCTGTTTCATGCTGTTAGATGATGCTGAAATGAGAC 1200
Db 1806 AGGAAGTAGTATTGGTAGTATTCTGTTTCATGCTGTTAGATGATGCTGAAATGAGAC 1865
QY 1201 CGCTTCATTTTGAATGCTGGGTTTCGTAGATGATTCACATGTTCAATACGGAATATCC 1260
Db 1866 CGCTTCATTTTGAATGCTGGGTTTCGTAGATGATTCACATGTTCAATACGGAATATCC 1925
QY 1261 TGAATCTCAGCTGCCCAACAGAGCTCCGACACAGCTAGAGCGCAAGACCCCTGG 1320
Db 1926 TGAATCTCAGCTGCCCAACAGAGCTCCGACACAGCTAGAGCGCAAGACCCCTGG 1985
QY 1321 AGATGACAGTGTCTGACGCTGGCAGATGCTCAGAAAGCTTTTGAAGCGGCTCTTAGG 1380

Db 1986 AGATGACAGTGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 2045
QY 1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTTAGACAGATCGCTCTTCTGCTGCTGT 1440
Db 2046 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTTAGACAGATCGCTCTTCTGCTGCTGT 2105
QY 1441 TGTGAGCGCAGGAGTA 1456
Db 2106 TGTGAGCGCAGGAGTA 2121
RESULT 3
US-09-564-479-1
; Sequence 1, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OOMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(2053)
US-09-564-479-1
Query Match 99.9%; Score 1455; DB 10; Length 2156;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATAAAACTTTTAAAAACAGGCTCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 60
Db 1 ATAAAACTTTTAAAAACAGGCTCGCAATTAATTAATTAATTAATTAATTAATTAATTTT 60
QY 61 ATAAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db 61 ATAAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
QY 121 AGGTCTCTATAGCAAGAAACAGAACCGCACACCTCCCGCAGATCTTTCTGCTCAAGGATTGA 180
Db 121 AGGTCTCTATAGCAAGAAACAGAACCGCACACCTCCCGCAGATCTTTCTGCTCAAGGATTGA 180
QY 181 GCGAGTGCAGCAATAAGAGTGCAGAGCTCAAGAGATAGAGAGTGCAGAGTGCAGAGTGCAGAG 240
Db 181 GCGAGTGCAGCAATAAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAG 240
QY 241 TAAAGAACTTAAGACCGATTCGTAGAGCGATGAGGAGCATCTTTCGCTCTGCAAGTGAATGC 300
Db 241 TAAAGAACTTAAGACCGATTCGTAGAGCGATGAGGAGCATCTTTCGCTCTGCAAGTGAATGC 300
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGTATTCCTTCTAGTAACAGCTCGCTCTTCTACTAG 360
Db 301 TCTCATGAGTCTGGCAGATAAGCTGGTATTCCTTCTAGTAACAGCTCGCTCTTCTACTAG 360
QY 361 CAGATCTGAGACGTGGAGTCAACGACGACCGCCTACGCTCTCCACCCACGTT 420
Db 361 CAGATCTGAGACGTGGAGTCAACGACGACCGCCTACGCTCTCCACCCACGTT 420
QY 421 TGAATGATTAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT 480

421 TGAATTAATAAGACTCAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACT 480
481 AGCTGACATACAGGCTGTTTGGTGAGGCTCAGAGTCTCTCACTAATAAAGGATAC 540
481 AGCTGACATACAGGCTGTTTGGTGAGGCTCAGAGTCTCTCACTAATAAAGGATAC 540
541 AGCGCTACTGATGAGGAAACCGCATCGCTCGGAGTGGGAAACTAAGATGCCGATGC 600
541 AGCGCTACTGATGAGGAAACCGCATCGCTCGGAGTGGGAAACTAAGATGCCGATGC 600
601 AGTTAAAGTTCGCGCGCAAAATACAGAATTAGCGAAATATGCTTCGGATAACCAAGCGAT 660
601 AGTTAAAGTTCGCGCGCAAAATACAGAATTAGCGAAATATGCTTCGGATAACCAAGCGAT 660
661 TCTTGACTCTTAGTAACTGACTCTCTTCCAGCTCTTACAGGCTCTTCTCCCAATC 720
661 TCTTGACTCTTAGTAACTGACTCTCTTCCAGCTCTTACAGGCTCTTCTCCCAATC 720
721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
781 CCCAGGAAACCGCTGCAATCTTCAATCTTTAGTTCAGACAGATGCTACAGCGAC 840
781 CCCAGGAAACCGCTGCAATCTTCAATCTTTAGTTCAGACAGATGCTACAGCGAC 840
841 ACAGATAGAGAAAGATGAAATGCAATGAGGATGATATTTTGAGGACAGACAGCGCTAG 900
841 ACAGATAGAGAAAGATGAAATGCAATGAGGATGATATTTTGAGGACAGACAGCGCTAG 900
901 TGGAGCTCTAGAAAATGCTAAATCTTAATAACAGTATAAGCAACAPAGATTGAGTAAAGC 960
901 TGGAGCTCTAGAAAATGCTAAATCTTAATAACAGTATAAGCAACAPAGATTGAGTAAAGC 960
961 AGCAATCGTACTGCTAGACACAATAAGCTGAGCTGAGCTGAGCAAAAGTTCCCGACTCC 1020
961 AGCAATCGTACTGCTAGACACAATAAGCTGAGCTGAGCTGAGCAAAAGTTCCCGACTCC 1020
1021 AATTTCTCAAGAACGGCAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCA 1080
1021 AATTTCTCAAGAACGGCAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCA 1080
1081 ACCTCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGAGGCTCCAAAGCAACA 1140
1081 ACCTCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGAGGCTCCAAAGCAACA 1140
1141 AGGAAGTAGTATGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC 1200
1141 AGGAAGTAGTATGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC 1200
1201 CGCTTCCATTTTGTCTGGGTTTCTGTCAGATGATTTCAATGTTCAATACGGAATAATCC 1260
1201 CGCTTCCATTTTGTCTGGGTTTCTGTCAGATGATTTCAATGTTCAATACGGAATAATCC 1260
1261 TGATTTCAAGCTGCGCAACAGAGCTCCGACGCAAGCTGAGCGCAAGAGCGGCTGG 1320
1261 TGATTTCAAGCTGCGCAACAGAGCTCCGACGCAAGCTGAGCGCAAGAGCGGCTGG 1320
1321 AGATGACAGTCTGCTGACGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
1321 AGATGACAGTCTGCTGACGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
1381 TAAAGCTGGCAACAACAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCT 1440
1381 TAAAGCTGGCAACAACAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCT 1440
1441 TGTGAGCGGAGGAGT 1455
1441 TGTGAGCGGAGGAGT 1455

US-10-289-762-1/c
Sequence 1, Application US/10289762
Publication No. US20040006218A1
GENERAL INFORMATION:
APPLICANT: Griffais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/10/289,762
CURRENT FILING DATE: 2003-03-27
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(15440)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (105001)..(120000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (120001)..(135000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (135001)..(150000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (150001)..(165000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (165001)..(180000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (180001)..(195000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (195001)..(210000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:

```
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (450001)..(465000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (510001)..(525000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (525001)..(540000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (540001)..(555000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (555001)..(570000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (570001)..(585000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (585001)..(600000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600001)..(615000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (615001)..(630000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (630001)..(645000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (645001)..(660000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (660001)..(675000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (675001)..(690000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
```


Query Match 99.2%; Score 1444; DB 15; Length 1230025;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 1 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTAGAGACTTTTTTTTTTTTTTTT 60
Db 828698 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTAGAGACTTTTTTTTTTTTTTTT 828699
QY 61 ATAAATAAATAAAGATTTTATTATTATTTTATGTTTAAATCCCTATTCGTCC 120
Db 828638 ATAAATAAATAAAGATTTTATTATTATTTTATGTTTAAATCCCTATTCGTCC 828579
QY 121 AGGTCTATAGAGAAACAGACGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA 180
Db 828578 AGGTCTATAGAGAAACAGACGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA 828519
QY 181 GCGAGTGCAGCAATAAGAGTGGGAGCTCAAGATAGCAGGTGCGGAGCTAAGCC 240
```

Db 828518 GCGAGTGCAGCAATAAGAGTGGGAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC 828459
Qy 241 TAAAGAACTAAGACCGAATCTGTAGAGCGATGGAGCATCTTGGCTTCTGCAGTGAATGC 300
Db 828458 TAAAGAACTAAGACCGAATCTGTAGAGCGATGGAGCATCTTGGCTTCTGCAGTGAATGC 828399
Qy 301 TCTCATAGTCTGGCAGATAAGCTGGTATTCTTCTAGTAAACAGCTGCTTCTTACTAG 360
Db 828398 TCTCATAGTCTGGCAGATAAGCTGGTATTCTTCTAGTAAACAGCTGCTTCTTACTAG 828339
Qy 361 CAGATCTCAGACGTGGAGCTCAACGACGAGCGGACCTAGCCCTCTCCACCCAGCTT 420
Db 828338 CAGATCTCAGACGTGGAGCTCAACGACGAGCGGACCTAGCCCTCTCCACCCAGCTT 828279
Qy 421 TGATGATTATAGACTCAAGCGCAACACAGCTTACGATCTATCTTTACTCAACATCACT 480
Db 828278 TGATGATTATAGACTCAAGCGCAACACAGCTTACGATCTATCTTTACTCAACATCACT 828219
Qy 481 AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAATATAAAGGATAC 540
Db 828218 AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAATATAAAGGATAC 828159
Qy 541 AGCGCTACTGATAGGAAACCGCAATCGCTCGGAGTGGAACTAAGATGCCGATGC 600
Db 828158 AGCGCTACTGATAGGAAACCGCAATCGCTCGGAGTGGAACTAAGATGCCGATGC 828099
Qy 601 AGTTAAAGTTGGCGCGCAATTTACAGAAATACAGAAATAGCTTCGGATTAACCAAGCGAT 660
Db 828098 AGTTAAAGTTGGCGCGCAATTTACAGAAATAGCTTCGGATTAACCAAGCGAT 828039
Qy 661 TCTTGACTCTTTAGGTAAATGACTGCTTCCTGACCTCTTACAGGCTGCTTCTTCAATC 720
Db 828038 TCTTGACTCTTTAGGTAAATGACTGCTTCCTGACCTCTTACAGGCTGCTTCTTCAATC 827979
Qy 721 TGTAGCAACATAACAGAGCTGAGCTTCTTAAAGAGATGCAAGTAAACCCAGTAGT 780
Db 827978 TGTAGCAACATAACAGAGCTGAGCTTCTTAAAGAGATGCAAGTAAACCCAGTAGT 827919
Qy 781 CCCAGGAAAACGCTGCAATTTGCTCAATCTTTAGTTAGTACAGACAGATGCTTACAGCGAC 840
Db 827918 CCCAGGAAAACGCTGCAATTTGCTCAATCTTTAGTTAGTACAGACAGATGCTTACAGCGAC 827859
Qy 841 ACAGATAGAGAAGATGGAAATCGGATAGGATGATATTTTGCAGGACAGAAACGCTAG 900
Db 827858 ACAGATAGAGAAGATGGAAATCGGATAGGATGATATTTTGCAGGACAGAAACGCTAG 827799
Qy 901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGATATAAGCAACATAGTTCAAGCTAAAGC 960
Db 827798 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGATATAAGCAACATAGTTCAAGCTAAAGC 827739
Qy 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 1020
Db 827738 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 827679
Qy 1021 AATTCTTCAAGAACGCGAACAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
Db 827678 AATTCTTCAAGAACGCGAACAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 827619
Qy 1081 ACCTGCAGATGTTCTGTGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 1140
Db 827618 ACCTGCAGATGTTCTGTGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 827559
Qy 1141 AGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTTTAGATGATGCTGAAAAATGAGAC 1200
Db 827558 AGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTTTAGATGATGCTGAAAAATGAGAC 827499
Qy 1201 CGCTTCCATTTGATGTTCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATAATCC 1260
Db 827498 CGCTTCCATTTGATGTTCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATAATCC 827439
Qy 1261 TGAATCTCAAGCTGCCCAACAGAGCTCGCAGCACAGCTAGCAGCGAGGCGAAGCGCTGG 1320

Db 827438 TGATTCTCAAGCTCCCAACAGGAGCTCGCAGCAAAAGCTAGACGAGCAAAAGCGCTGG 827379
Qy 1321 AGATGACAGTCTGCTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGG 1380
Db 827378 AGATGACAGTCTGCTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGG 827319
Qy 1381 TAAAGCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTG- 1439
Db 827318 TAAAGCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTG- 827259
Qy 1440 TTGCTAGCGCAGGAGT 1455
Db 827258 TTGCTAGCGCAGGAGT 827243

RESULT 5
US-09-841-132-385
; Sequence 385, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yaser A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469CB
; CURRENT APPLICATION NUMBER: US/09/841132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 385
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-385



Query Match 93.1%; Score 1355; DB 9; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ATGGTTAATCTTATGGTCCAGGTCCTATAGCAAGAAACAGAACGACACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTTATGGTCCAGGTCCTATAGCAAGAAACAGAACGACACACCTCCCGCAGAT 60
Qy 161 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATTAAGAGTGGGAAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATTAAGAGTGGGAAGCTCAAGAATA 120
Qy 221 GCAGTGGGGAAGCTTAAGCTTAAAGAACTTAAGACCGATTTCTAGAGCGATGGGCAATC 280
Db 121 GCAGTGGGGAAGCTTAAAGCTTAAAGAACTTAAGACCGATTTCTAGAGCGATGGGCAATC 180
Qy 281 TTGCGTTCTGCAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGT 340
Db 181 TTGCGTTCTGCAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGT 240
Qy 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGACCGCACT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGACCGCACT 300
Qy 401 AGCGCTCTCCACCCAGCTTTGATGATTATAGACTCAAGCGCAAAACAGCTTACGATACT 460
Db 301 AGCGCTCTCCACCCAGCTTTGATGATTATAGACTCAAGCGCAAAACAGCTTACGATACT 360
Qy 461 ATCTTTACTCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGGCTCCAGGATGCT 520
Db 361 ATCTTTACTCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGGCTCCAGGATGCT 420
Qy 521 GTCACTAATATAAAGGATACAGCGGCTACTGATGAGGAAACCGCAATGCTGCGGAGTGG 580
Db 421 GTCACTAATATAAAGGATACAGCGGCTACTGATGAGGAAACCGCAATGCTGCGGAGTGG 480
Qy 581 GAACTAAGATGCCGATGCTAGTTAAGTTGGCGGCAAAATTAAGAATTAGCGAATAT 640

481 GAACTAAGAAATCCGATGAGTTAAAGTTGGCGCGCAAAATTAAGAAATAGCGAAATAT 540
641 GCTTCGGATAACCAAGCGATCTTTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 700
541 GCTTCGGATAACCAAGCGATCTTTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 600
701 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAAAGAGAGCTGAGCTTCTTAAGAG 760
601 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAAAGAGAGCTGAGCTTCTTAAGAG 660
761 ATGCAAGATTAACCCAGTAGTCCAGGAGAAACCGCTGCAATCTCAATCTTCTAGTTGAT 820
661 ATGCAAGATTAACCCAGTAGTCCAGGAGAAACCGCTGCAATCTCAATCTTCTAGTTGAT 720
821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 880
721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 780
881 TTTGCGAGGACAGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAAGCTATAAGC 940
781 TTTGCGAGGACAGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAAGCTATAAGC 840
941 AACATAGATTGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
841 AACATAGATTGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900
1001 AAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 1060
901 AAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 960
1061 AAGATCTTAAATAATCAAACTGAGAGCTGCTGATGTTCCAAATCCAGGAACTACA 1120
961 AAGATCTTAAATAATCAAACTGAGAGCTGCTGATGTTCCAAATCCAGGAACTACA 1020
1121 GTTGGAGGCTCCAGCAACCAAGAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTA 1180
1021 GTTGGAGGCTCCAGCAACCAAGAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTA 1080
1181 GATGATGCTGAAATAGAGCGCTTCCATTTGATGCTGGGTTTCTGAGATGATTCAC 1240
1081 GATGATGCTGAAATAGAGCGCTTCCATTTGATGCTGGGTTTCTGAGATGATTCAC 1140
1241 ATGTTCAATACGGAATCTGATTTCTCAAGCTGCTCCAAAGAGAGCTGCGAGCAGCAGCT 1300
1141 ATGTTCAATACGGAATCTGATTTCTCAAGCTGCTCCAAAGAGAGCTGCGAGCAGCAGCT 1200
1301 AGAGCAGGAAAGCGCTGGAGATGACAGTCTGCTGCTGAGGCTGGCAGATGCTCAGAAA 1360
1201 AGAGCAGGAAAGCGCTGGAGATGACAGTCTGCTGCTGAGGCTGGCAGATGCTCAGAAA 1260
1361 GCTTTAGAGCGGCTPAGTAAAGCTGGGCAACAAAGAGGCTGCTCAATGCTTTAGGA 1420
1261 GCTTTAGAGCGGCTPAGTAAAGCTGGGCAACAAAGAGGCTGCTCAATGCTTTAGGA 1320
1421 CAGATGCTTCTGCTGCTGTTGAGCGCAGGAGT 1455
1321 CAGATGCTTCTGCTGCTGTTGAGCGCAGGAGT 1355

RESULT 6
US-09-841-260-63
; Sequence 63, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-63

Query Match 93.1%; Score 1355; DB 10; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGCTTTAATCTTATTCGTCAGGTCTCTATAGCAAAACAGACGACACCTCCGCGAGAT 160
Db 1 ATGCTTTAATCTTATTCGTCAGGTCTCTATAGCAAAACAGACGACACCTCCGCGAGAT 60
QY 161 CTTTCTGCTCAAGGATTTGGAGCGAGTGGAGCAAAATTAAGAGTGGGAAAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTTGGAGCGAGTGGAGCAAAATTAAGAGTGGGAAAGCTCAAGAATA 120
QY 221 GCAGGTGCGGAAAGCTTAAGCTTAAAGAAATCTAAGACCGATTTCTGTAGAGCGATCGACATC 280
Db 121 GCAGGTGCGGAAAGCTTAAGCTTAAAGAAATCTAAGACCGATTTCTGTAGAGCGATCGACATC 180
QY 281 TTGCGTTCTCGAGTGAATGCTCTCATGACTCTGCGAGATAAGCTGGTATTGCTTCTAGT 340
Db 181 TTGCGTTCTCGAGTGAATGCTCTCATGACTCTGCGAGATAAGCTGGTATTGCTTCTAGT 240
QY 341 AACAGCTCGTCTTCTACTAGCAGATCTGCGAGCTGGAATCAACGACGACGCGCACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCGAGCTGGAATCAACGACGACGCGCACCT 300
QY 401 AGCCTCTCTCCAGCCACGTTTGAATTAAGACTCAAGCGCAAAACAGCTTACGATAT 460
Db 301 AGCCTCTCTCCAGCCACGTTTGAATTAAGACTCAAGCGCAAAACAGCTTACGATAT 360
QY 461 ATCTTTACCTCAACATCAGTCTGATGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCT 520
Db 361 ATCTTTACCTCAACATCAGTCTGATGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCT 420
QY 521 GTCACATAATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 580
Db 421 GTCACATAATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 480
QY 581 GAAACTAAGATGCCAGTGAATTAAGTTGGCGCGCAAAATTAAGAAATTAAGGAAATAT 640
Db 481 GAAACTAAGATGCCAGTGAATTAAGTTGGCGCGCAAAATTAAGAAATTAAGGAAATAT 540
QY 641 GCTTCGGATAAACCAAGCGATTTCTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 700
Db 541 GCTTCGGATAAACCAAGCGATTTCTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAAAGAGAGCTGAGCTTCTTAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAAAGAGAGCTGAGCTTCTTAAGAG 660
QY 761 ATGCAAGATAACCCAGTAGTCCAGGAGAAACCGCTGCAATCTCAATCTTCTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTAGTCCAGGAGAAACCGCTGCAATCTCAATCTTCTAGTTGAT 720
QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 880
Db 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 780
QY 881 TTTGCGAGGACAGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAAGCTATAAGC 940
Db 781 TTTGCGAGGACAGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAAGCTATAAGC 840
QY 941 AACATAGATTGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900
QY 1001 AAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 1060

Db 901 AAAAAGTTCCCGACTCTCCAAATCTTTCAAGAAGCGGAACAAATGGTAATACAGGCTGAG 960
Qy 1061 AAAGATCTTAAATAATCAAAACCTGCAGATGGTTCTGATGTTCCAAATPCCAGGAACCTACA 1120
Db 961 AAAGATCTTAAATAATCAAAACCTGCAGATGGTTCTGATGTTCCAAATPCCAGGAACCTACA 1020
Qy 1121 GTTGGAGCTCCAAAGCAACAAAGGAAGTAGTATGGTAGTATGCTGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGCTCCAAAGCAACAAAGGAAGTAGTATGGTAGTATGCTGTTTCCATGCTGTTA 1080
Qy 1181 GATGATCTGAAATGAGACCGCTTCCATTTGATGTTCTGGTTTCGTCAGATGATTCAC 1240
Db 1081 GATGATCTGAAATGAGACCGCTTCCATTTGATGTTCTGGTTTCGTCAGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACACAAGCT 1300
Db 1141 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACACAAGCT 1200
Qy 1301 AGACGAGGAAGCGCTGGAGATGACAGTCTGTCGAGCGCTGGCGAGATGCTCAGAAA 1360
Db 1201 AGACGAGGAAGCGCTGGAGATGACAGTCTGTCGAGCGCTGGCGAGATGCTCAGAAA 1260
Qy 1361 GCTTTAGAGCGGCTCTAGGTAAAGCTGGGCAACAAACAGGCGATCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGGCTCTAGGTAAAGCTGGGCAACAAACAGGCGATCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1455
Db 1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1355

RESULT 7
US-10-007-693-63
; Sequence 63, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.51-5C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-63

Query Match 93.1%; Score 1355; DB 13; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ATGGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGCACACCTCCGCGAGAT 160
Db 1 ATGGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGCACACCTCCGCGAGAT 60
Qy 161 CTTTCTGCTCAAGGATTTGGAGCGAGTGCAGCAATAAAGAGTGGCGAAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTTGGAGCGAGTGCAGCAATAAAGAGTGGCGAAGCTCAAGAATA 120
Qy 221 GCAGTGGGGAAGCTTAAGCTTAAGAACTTAAGCCGATTTGTAGAGCGATGGAGCATC 280
Db 121 GCAGTGGGGAAGCTTAAGCTTAAGAACTTAAGCCGATTTGTAGAGCGATGGAGCATC 180
Qy 281 TTGCGTTCGAGTCAATGCTCTCATGAGTCTGGCAGATAGCTGGGTATGCTTCTAGT 340
Db 181 TTGCGTTCGAGTCAATGCTCTCATGAGTCTGGCAGATAGCTGGGTATGCTTCTAGT 240
Qy 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGAGTCAACGACAGCGACCCACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGAGTCAACGACAGCGACCCACCT 300

Qy 401 ACGCCTCTCCACCCAGCTTTGATGATTATAGACTCAAGCGCAACAGCTTACGATPACT 460
Db 301 ACGCCTCTCCACCCAGCTTTGATGATTATAGACTCAAGCGCAACAGCTTACGATPACT 360
Qy 461 ATCTTTACCTCAACATCAGTCTAGTGCATACAGGCTGCTTTGGTGGCTTCCAGGATGCT 520
Db 361 ATCTTTACCTCAACATCAGTCTAGTGCATACAGGCTGCTTTGGTGGCTTCCAGGATGCT 420
Qy 521 GTCACCTAAATAAAGGATACAGCGGCTTACTGATGAGGAAACCCGCAATCGCTGGGAGTGG 580
Db 421 GTCACCTAAATAAAGGATACAGCGGCTTACTGATGAGGAAACCCGCAATCGCTGGGAGTGG 480
Qy 581 GAACTAAGAAATGCGCGATGCGATTAAAGTTGGCGCGCAAAATTCAGAAATAGCGAAATAT 640
Db 481 GAACTAAGAAATGCGCGATGCGATTAAAGTTGGCGCGCAAAATTCAGAAATAGCGAAATAT 540
Qy 641 GCTTCGGATTAACCAAGGATTTCTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 700
Db 541 GCTTCGGATTAACCAAGGATTTCTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 600
Qy 701 CAGGCTCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAG 760
Db 601 CAGGCTCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAG 660
Qy 761 ATGCAAGATTAACCCAGTGTCCAGGGAACCGCTGCAATTCGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATTAACCCAGTGTCCAGGGAACCGCTGCAATTCGCTCAATCTTTAGTTGAT 720
Qy 821 CAGACAGATCTACAGCGACACAGATAGAGAAAGTGGAAATCGATTAGGGATGCATAT 880
Db 721 CAGACAGATCTACAGCGACACAGATAGAGAAAGTGGAAATCGATTAGGGATGCATAT 780
Qy 881 TTTGCGAGCAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAGC 940
Db 781 TTTGCGAGCAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAGC 840
Qy 941 AACATAGTTTCCAGTAAAGCAGCAATCGCTACTGCTNAGACACAAATAGCTGAGCTCAG 1000
Db 841 AACATAGTTTCCAGTAAAGCAGCAATCGCTACTGCTNAGACACAAATAGCTGAGCTCAG 900
Qy 1001 AAAAAGTTCCCGACTCTCCAAATCTTCAAGAAGCGGAACAAATGGTAATACAGGCTGAG 1060
Db 901 AAAAAGTTCCCGACTCTCCAAATCTTCAAGAAGCGGAACAAATGGTAATACAGGCTGAG 960
Qy 1061 AAAGATCTTAAATAATCAAAACCTGCAGATGGTTCTGATGTTCCAAATPCCAGGAACCTACA 1120
Db 961 AAAGATCTTAAATAATCAAAACCTGCAGATGGTTCTGATGTTCCAAATPCCAGGAACCTACA 1020
Qy 1121 GTTGGAGCTCCAAAGCAACAAAGGAAGTAGTATGGTAGTATGCTGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGCTCCAAAGCAACAAAGGAAGTAGTATGGTAGTATGCTGTTTCCATGCTGTTA 1080
Qy 1181 GATGATCTGAAATGAGACCGCTTCCATTTGATGTTCTGGTTTCGTCAGATGATTCAC 1240
Db 1081 GATGATCTGAAATGAGACCGCTTCCATTTGATGTTCTGGTTTCGTCAGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACACAAGCT 1300
Db 1141 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACACAAGCT 1200
Qy 1301 AGACGAGGAAGCGCTGGAGATGACAGTCTGCTGAGCGCTGGCGAGATGCTCAGAAA 1360
Db 1201 AGACGAGGAAGCGCTGGAGATGACAGTCTGCTGAGCGCTGGCGAGATGCTCAGAAA 1260
Qy 1361 GCTTTAGAGCGGCTCTAGGTAAAGCTGGGCAACAAACAGGCGATCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGGCTCTAGGTAAAGCTGGGCAACAAACAGGCGATCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1455
Db 1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1355

RESULT 8
US-10-312-273-14
; Sequence 14, Application US/10312273
; Publication No. US2004005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035W0
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 14
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-14

Query Match 93.1%; Score 1355; DB 15; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 ATGGTTAATCCTATTGGTCCAGGCTCTATAGACGAAACAGACGACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCCTATTGGTCCAGGCTCTATAGACGAAACAGACGACACCTCCCGCAGAT 60
QY 161 CTTTCTGCTCAAGGATTGGAGGAGTGCAGCAAAATAGAGTGGGAGCTCAAGAAATA 220
Db 61 CTTTCTGCTCAAGGATTGGAGGAGTGCAGCAAAATAGAGTGGGAGCTCAAGAAATA 120
QY 221 GCAGGTGCGGAAGCTAAGCCTAAAGAAATCTAAGACCGATTCTGTAGAGCGATGGAGCATC 280
Db 121 GCAGGTGCGGAAGCTAAGCCTAAAGAAATCTAAGACCGATTCTGTAGAGCGATGGAGCATC 180
QY 281 TTGCGTTCTGCAAGTAAATGCTCTCATAGTCTTGGAGATAAGCTGGGTATTGCTTCTAGT 340
Db 181 TTGCGTTCTGCAAGTAAATGCTCTCATAGTCTTGGAGATAAGCTGGGTATTGCTTCTAGT 240
QY 341 AACAGTCTGCTTCTACTACGATCTGCAGAGTGCAGCTCAACGACGACGACCGCACCT 400
Db 241 AACAGTCTGCTTCTACTACGATCTGCAGAGTGCAGCTCAACGACGACGACCGCACCT 300
QY 401 ACGCCTCTCCACCCAGCTTTGATGATTATAAGACTCAAGCGCAACAGCTTACGATACT 460
Db 301 ACGCCTCTCCACCCAGCTTTGATGATTATAAGACTCAAGCGCAACAGCTTACGATACT 360
QY 461 ATCTTACCTCAACATCACTAGTGCATACAGGCTGCTTTGGTGGCCCTCCAGGATGCT 520
Db 361 ATCTTACCTCAACATCACTAGTGCATACAGGCTGCTTTGGTGGCCCTCCAGGATGCT 420
QY 521 GTCACTAATAAAGATACAGGCTTACTGATGAGGAAACCCCAATCGCTGCGGAGTGG 580
Db 421 GTCACTAATAAAGATACAGGCTTACTGATGAGGAAACCCCAATCGCTGCGGAGTGG 480
QY 581 GAAACTAAGAAATCCCGATGAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGGGAATAT 640
Db 640 GAAACTAAGAAATCCCGATGAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGGGAATAT 640

Db 481 GAAACTAAGAAATCCCGATGAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGGGAATAT 540
QY 641 GCTTCGATAAACCAAGCGGATCTTTGACTCTTTAGGTAAACTGACTTCTCCGACCTCTTA 700
Db 541 GCTTCGATAAACCAAGCGGATCTTTGACTCTTTAGGTAAACTGACTTCTCCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTTCTTAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTTCTTAAGAG 660
QY 761 ATGCAAGATAACCCAGTACTGCCAGGGAAGACCCCTGCAATTCCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTACTGCCAGGGAAGACCCCTGCAATTCCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATGCTACACGCGACACAGATAGAGAAAGATGGAAATCGGATTAGGGATGCATAT 880
Db 721 CAGACAGATGCTACACGCGACACAGATAGAGAAAGATGGAAATCGGATTAGGGATGCATAT 780
QY 881 TTTCAGAGCAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGC 940
Db 781 TTTCAGAGCAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGC 840
QY 941 AACATAGATTCAAGTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTCAAGTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900
QY 1001 AAAAAGTTCCCGGACTCTCCAAATTTCTCAAGAGCGGAACAAATGGTAAATACAGGCTGAG 1060
Db 901 AAAAAGTTCCCGGACTCTCCAAATTTCTCAAGAGCGGAACAAATGGTAAATACAGGCTGAG 960
QY 1061 AAAGATCTTAAATAATCAAACTGCAGATGTTCTGATGTTCCAAATCCAGGAATCTACA 1120
Db 961 AAAGATCTTAAATAATCAAACTGCAGATGTTCTGATGTTCCAAATCCAGGAATCTACA 1020
QY 1121 GTTGAGGCTCCAAGCAACAGGAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTA 1180
Db 1021 GTTGAGGCTCCAAGCAACAGGAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTA 1080
QY 1181 GATGATGCTGAAATAGAGACCGCTTCCATTTGATGTTGCGTTCGTCAGATGATTCAC 1240
Db 1081 GATGATGCTGAAATAGAGACCGCTTCCATTTGATGTTGCGTTCGTCAGATGATTCAC 1140
QY 1241 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCACAGGAGCTCCGACAGAGCT 1300
Db 1141 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCACAGGAGCTCCGACAGAGCT 1200
QY 1301 AGAGCAGCAAAAGCCGCTGGAGATGACAGTGTGCTGCGAGCGTGGCAGATGCTCAGAAA 1360
Db 1201 AGAGCAGCAAAAGCCGCTGGAGATGACAGTGTGCTGCGAGCGTGGCAGATGCTCAGAAA 1260
QY 1361 GCTTAGAAGCGGCTCTAGGTTAAAGCTGGGCAACAAAGGGGATCTCAATGCTTTAGGA 1420
Db 1261 GCTTAGAAGCGGCTCTAGGTTAAAGCTGGGCAACAAAGGGGATCTCAATGCTTTAGGA 1320
QY 1421 CAGATCGCTTCTGCTGCTGTTGTCGCGCAGAGT 1455
Db 1321 CAGATCGCTTCTGCTGCTGTTGTCGCGCAGAGT 1355

RESULT 9

US-09-564-479-3
; Sequence 3, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OOMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03

09/09/00


```
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1749)
US-09-564-479-3

Query Match
 79.1%; Score 1151; DB 10; Length 1852;
Best Local Similarity 100.0%; Pred. No. 3.2e-289;
Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 364
Db 1 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
QY 365 TCTGCAGAGCTGAGCTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTTTGTAT 424
Db 61 TCTGCAGAGCTGAGCTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTTTGTAT 120
QY 425 GATTATAAGACTCAAGCGCAACAGCTTACGATCTATCTTTTACCTCAACATCACTAGCT 484
Db 121 GATTATAAGACTCAAGCGCAACAGCTTACGATCTATCTTTTACCTCAACATCACTAGCT 180
QY 485 GACATACAGGCTGCTTGGTGAAGCTCCAGGATGCTGTCACCTAATAATAAGGATACAGG 544
Db 181 GACATACAGGCTGCTTGGTGAAGCTCCAGGATGCTGTCACCTAATAATAAGGATACAGG 240
QY 545 GCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAAACCTAAGAACTCCGATCGAGTT 604
Db 241 GCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAAACCTAAGAACTCCGATCGAGTT 300
QY 605 AAAGTTGGCGCCAAATTACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 664
Db 301 AAAGTTGGCGCCAAATTACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 360
QY 665 GACTCTTTAGTAACTGACTTCTTGGACCTTACAGGCTGCTTCTTCCAACTCTGTA 724
Db 361 GACTCTTTAGTAACTGACTTCTTGGACCTTACAGGCTGCTTCTTCCAACTCTGTA 420
QY 725 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATCAAGATAACCCAGTAGTCCCA 784
Db 421 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATCAAGATAACCCAGTAGTCCCA 480
QY 785 GGGAAACCCCTGGAATGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCAGCAG 844
Db 481 GGGAAACCCCTGGAATGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCAGCAG 540
QY 845 ATAGAGAAGATGGAAATCGGATAGGAGTGCATATTTTGCAGGACAGAACTAGTGA 904
Db 541 ATAGAGAAGATGGAAATCGGATAGGAGTGCATATTTTGCAGGACAGAACTAGTGA 600
QY 905 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGCAGCA 964
Db 601 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGCAGCA 660
QY 965 ATCGCTACTGCTAAGCACAATAATAGCTGAAGCTCAGAAAAGTCCCGACTCTCCAAT 1024
Db 661 ATCGCTACTGCTAAGCACAATAATAGCTGAAGCTCAGAAAAGTCCCGACTCTCCAAT 720
QY 1025 CTTCAAGAGCGGAAACAAATGTTAATACAGGCTGAGAAAGATCTTAAATAATCAAACT 1084
Db 721 CTTCAAGAGCGGAAACAAATGTTAATACAGGCTGAGAAAGATCTTAAATAATCAAACT 780
QY 1085 GCAGATGGTCTGATGTTCCAAATCCAGGAATCTACAGTTGGAGGCTCCCAAGCAACAGGA 1144
Db 1144 GCAGATGGTCTGATGTTCCAAATCCAGGAATCTACAGTTGGAGGCTCCCAAGCAACAGGA 840
QY 1145 AGTAGTATTGGTAGTATTTCGTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT 1204
Db 841 AGTAGTATTGGTAGTATTTCGTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT 900
QY 1205 TCCATTTTGTGCTGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATAATCCTGAT 1264
Db 901 TCCATTTTGTGCTGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATAATCCTGAT 960
QY 1265 TCTCAAGCTGCCCAACAGGAGCTCGCAGCACAAGCTAGACAGCGAAGCGCTGGAGAT 1324
Db 961 TCTCAAGCTGCCCAACAGGAGCTCGCAGCACAAGCTAGACAGCGAAGCGCTGGAGAT 1020
QY 1325 GACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGGCTCTAGGTAAA 1384
Db 1021 GACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGGCTCTAGGTAAA 1080
QY 1385 GCTGGGCAACACAGGCGCATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTTGTG 1444
Db 1081 GCTGGGCAACACAGGCGCATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTTGTG 1140
QY 1445 AGCGCAGGAGT 1455
Db 1141 AGCGCAGGAGT 1151

RESULT 10
US-09-905-119-1
; Sequence 1, Application US/09905119
; Patent No. US2002091096A1
; GENERAL INFORMATION:
; APPLICANT: Murdin, Andrew D
; APPLICANT: Dunn, Pamela L
; TITLE OF INVENTION: Chlamydia Antigens and Corresponding DNA Fragments and
; FILE REFERENCE: 19721-010 DIV
; CURRENT APPLICATION NUMBER: US/09/905,119
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/106,037
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/154,658
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/427,501
; PRIOR FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1366)
US-09-905-119-1

Query Match
 11.7%; Score 170; DB 9; Length 1550;
Best Local Similarity 100.0%; Pred. No. 2e-33;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTTAAACACAGGCTCGCANTTAATTATTAGTGAGAGCTTTTTTTTATTTTTT 60
Db 1381 ATAAATCTTTTAAACACAGGCTCGCANTTAATTATTAGTGAGAGCTTTTTTTTATTTTTT 1440
QY 61 ATAATAAACTAAAGATTTTTTATTTTTTTAGTTTTTATGTTTAAATCCTATTGTGTC 120
Db 1441 ATAATAAACTAAAGATTTTTTATTTTTTTAGTTTTTATGTTTAAATCCTATTGTGTC 1500
QY 121 AGGTCTTATAGCAAGAAACAGAACCGCACCTCCCGCAGATCTTCTGCTC 170
Db 1501 AGGTCTTATAGCAAGAAACAGAACCGCACCTCCCGCAGATCTTCTGCTC 1550
```



```

; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 13
; LENGTH: 1171
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-841-260-13

Query Match 3.3%; Score 48.6; DB 10; Length 1171;
Best Local Similarity 53.7%; Pred. No. 0.076;
Matches 124; Conservative 0; Mismatches 104; Indels 3; Gaps 1;

Qy 1167 TTTCATGCTGTTAGATGATGCTGAAATAAGAGACCGCTTCATTTTGATGCTCTGGGTTTC 1226
Db 207 TTTCCTTGTTGCTTGATGATGATAGACAATGAAATGGCAGCGATTCACCTGCAAGTTTTC 266

Qy 1227 GTCAGATGATTCACATGTTCAATACGGAATAATCCTGATTTCAAGCTGCCCAACAGGAGC 1286
Db 267 GATCTATGATCGAACCAATTTAATGTAACAATCTGCAACAGCTAAAGAGCTACAAGCTA 326

Qy 1287 TCGCAGCACAAAGCTAGACGACGCAAGCCGCTGGAGATGACACTGCTGCGACGCGTGG 1346
Db 327 TGGAGGCTCAGCTGA---CTGCCGATGTCAGATCAACTGTTGTTGGGATGGCGAGCTCC 383

Qy 1347 CAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGGTAAAGCTGGCGCAACAAC 1397
Db 384 CAGCCGAATACAGCAATCAAGATGCTCTTGGCAGAGCTTTGAAACAAC 434

RESULT 15
US-10-007-693-13
; Sequence 13, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 13
; LENGTH: 1171
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-10-007-693-13

Query Match 3.3%; Score 48.6; DB 13; Length 1171;
Best Local Similarity 53.7%; Pred. No. 0.076;
Matches 124; Conservative 0; Mismatches 104; Indels 3; Gaps 1;

Qy 1167 TTTCATGCTGTTAGATGATGCTGAAATAAGAGACCGCTTCATTTTGATGCTCTGGGTTTC 1226
Db 207 TTTCCTTGTTGCTTGATGATGATAGACAATGAAATGGCAGCGATTCACCTGCAAGTTTTC 266

Qy 1227 GTCAGATGATTCACATGTTCAATACGGAATAATCCTGATTTCAAGCTGCCCAACAGGAGC 1286
Db 267 GATCTATGATCGAACCAATTTAATGTAACAATCTGCAACAGCTAAAGAGCTACAAGCTA 326

Qy 1287 TCGCAGCACAAAGCTAGACGACGCAAGCCGCTGGAGATGACACTGCTGCGACGCGTGG 1346
Db 327 TGGAGGCTCAGCTGA---CTGCCGATGTCAGATCAACTGTTGTTGGGATGGCGAGCTCC 383

Qy 1347 CAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGGTAAAGCTGGCGCAACAAC 1397
Db 384 CAGCCGAATACAGCAATCAAGATGCTCTTGGCAGAGCTTTGAAACAAC 434

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:15 ; Search time 5658.52 Seconds
(without alignments)
11152.641 Million cell updates/sec

Title: US-10-608-559-5
Perfect score: 1456
Sequence: 1 aataatctttaaacaagg.....ctgttgagcaggaagta 1456

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_cv:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_man:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1456	100.0	1456	6	AX045135 Sequence
2	1456	100.0	2238	6	AX045137 Sequence
3	1455	99.9	2156	6	AX045131 Sequence
4	1455	99.9	10421	1	AE001654 Chlamydia
5	1455	99.9	12548	1	AE002165 Chlamydia
6	1455	99.9	300566	1	AE017159 Chlamydia
7	1455	99.9	300550	1	AP002547 Chlamydia
8	1444	99.2	2148	1	CHT76KDA
9	1444	99.2	110000	6	AR310754_08
10	1355	93.1	1956	6	AX338293 Sequence
11	1355	93.1	1956	6	AX349491 Sequence
12	1355	93.1	1956	6	AX361992 Sequence
13	1151	79.1	1852	6	AX045133 Sequence
14	204.2	14.0	304769	1	AE016994 Chlamydia
15	170	11.7	1550	6	BD243971 Chlamydia
16	170	11.7	1550	6	AR212731 Sequence
17	170	11.7	1550	6	AR428600 Sequence
18	148.4	10.2	150	6	ARI44067 Sequence
19	55.4	3.8	2000	6	AX655393 Sequence
20	54.4	3.7	1941	6	AX362019 Sequence
21	54.4	3.7	1944	6	AX786284 Sequence
22	54.4	3.7	7218	6	I66494 Sequence 14
23	54.4	3.7	10954	1	AE001333 Chlamydia
24	51.6	3.5	2205	6	AR317736 Sequence
25	51.4	3.5	1537	6	BD264444 Compounds
26	51.4	3.5	1537	6	AR229257 Sequence
27	51.4	3.5	1537	6	AR308168 Sequence
28	51.4	3.5	1537	6	AR321875 Sequence
29	51.4	3.5	1537	6	AX155814 Sequence
30	51.4	3.5	1537	6	AX361664 Sequence
31	51.4	3.5	154160	2	BX571709 Danio rerio
32	50.6	3.5	423	1	MC049
33	50.2	3.4	14698	1	AE002357 Chlamydia
34	49.4	3.4	5098	1	MCU51235 Mycoplasma
35	49	3.4	1990	6	AR268644 Sequence
36	49	3.4	2085	6	AR008322 Sequence
37	49	3.4	2085	6	AR027306 Sequence
38	49	3.4	2085	6	AR035771 Sequence
39	49	3.4	2085	6	AR037631 Sequence
40	49	3.4	2085	6	AR078816 Sequence
41	49	3.4	2085	6	AR079063 Sequence
42	49	3.4	2085	6	AR084296 Sequence
43	49	3.4	2085	6	I16740 Sequence 1
44	49	3.4	2085	6	I70472 Sequence 1
45	49	3.4	2085	6	I92332 Sequence 1

ALIGNMENTS

RESULT 1	AX045135	Sequence 5 from Patent WO0066739.	1456 bp	DNA	linear	PAT 24-NOV-2000
LOCUS	AX045135	Sequence 5 from Patent WO0066739.				
DEFINITION	AX045135					
ACCESSION	AX045135.1	GI:11343734				
VERSION	AX045135.1					
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						

Aventis Pasteur Limited (CA)	
Location/Qualifiers	
1. .1456	
/organism="Chlamydomophila pneumoniae"	
/mol_type="unassigned DNA"	
/db_xref="taxon:83558"	
101. .>1452	
CDS	
/note="unassigned protein product"	
/codon_start=1	
/transl_table=11	
/protein_id="CACL7379.1"	
/db_xref="GI:11343735"	
/db_xref="REMBL:CACL7379"	
/translation="MNPIGPGFIDETERTPPADLSAQGLEASRAANKSAEARIAGAE AKPKSKTSVERKILSSAVNALMSLADKLGIASSNSSSSTSRASVDSTTATATPT PPPTFDYKTQATYDTFTSTSLADIQAALVSLQDAVNI KDTAAEDETATIAAEW ETKNADAVKVAQIITELAKYASDNOAILDSGLKLSFDLLOAALLQSVANNKAAELL KEMONPVVPGKTPALQSLVDQTDATQIEKDGNALRDYAFQGNASGAVENAKSN NSISNIDSAKALIAKQIAEAKKFEDPSILQAEQVIOREKDLKNIKRPADGSDV PNPFTVVGSKKQSGSISGIRVSMLLDDDAENETASILMSGFQFMHMENTPDSQA QOELAAQFAAKAAGSDSAAALADAQKALEAALGKAGQOQOQILNALQGIASAAVSA G"	
ORIGIN	
Query Match	100.0%; Score 1456; DB 6; Length 1456;
Best Local Similarity	100.0%; Pred. No. 2.4e-301; Mismatches 0; Indels 0; Gaps 0;
Matches 1456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1	ATATAATCTTTAAAAACAGCTCGCATTAATATTAGTCAGAGCTTTTTTTTATTTTTT 60
DB 1	ATATAATCTTTAAAAACAGCTCGCATTAATATTAGTCAGAGCTTTTTTTTATTTTTT 60
QY 61	ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
DB 61	ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
QY 121	AGTCTCTATAGAGCAACAGAGCGACACCTCCGCGAGATCTTCTGCTCAGAGATTGGA 180
DB 121	AGTCTCTATAGAGCAACAGAGCGACACCTCCGCGAGATCTTCTGCTCAGAGATTGGA 180
QY 181	GGCGAGTCAGCAAAATAAGAGTCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC 240
DB 181	GGCGAGTCAGCAAAATAAGAGTCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC 240
QY 241	TAAAGAACTAAGACCGATTTCTGTAGAGCGATGGAGCATCTTCGCTTCTGAGTGAATGC 300
DB 241	TAAAGAACTAAGACCGATTTCTGTAGAGCGATGGAGCATCTTCGCTTCTGAGTGAATGC 300
QY 301	TCTCATGCTCTGGCAGATAAGCTGGGTATTCCTTCTAGTAAACAGCTGCTTCTACTAG 360
DB 301	TCTCATGCTCTGGCAGATAAGCTGGGTATTCCTTCTAGTAAACAGCTGCTTCTACTAG 360
QY 361	CAGATCTGAGACGTTGGATCAACGACGACCGCACCTACGCTCTCCACCCACGTT 420
DB 361	CAGATCTGAGACGTTGGATCAACGACGACCGCACCTACGCTCTCCACCCACGTT 420
QY 421	TGATGATTAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT 480
DB 421	TGATGATTAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT 480
QY 481	AGCTGACATACAGGCTGCTTTGGTGAAGCTTCAGGATGCTGTCACCTAATATAAGGATAC 540
DB 481	AGCTGACATACAGGCTGCTTTGGTGAAGCTTCAGGATGCTGTCACCTAATATAAGGATAC 540
QY 541	AGCGGCTACTGATGAGGAAACCGCAATCGTCGGAGTGGAAACTAAGAATGCCGATGC 600
DB 541	AGCGGCTACTGATGAGGAAACCGCAATCGTCGGAGTGGAAACTAAGAATGCCGATGC 600
QY 601	AGTTAAAGTTGGCGCCCAATACAGAAATAGCGAAATATGCTTCGGATAACAGACGAT 660
DB 601	AGTTAAAGTTGGCGCCCAATACAGAAATAGCGAAATATGCTTCGGATAACAGACGAT 660
QY 661	TCTTGACTCTTTAGGTAACCTGACTTCTTCGACCTCTTACAGGCTGCTTCTTCCAAATC 720
DB 661	TCTTGACTCTTTAGGTAACCTGACTTCTTCGACCTCTTACAGGCTGCTTCTTCCAAATC 720
QY 721	TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
DB 721	TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
QY 781	CCCAGGAAAAACGCTCGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC 840
DB 781	CCCAGGAAAAACGCTCGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC 840
QY 841	ACAGATAGAGAAAGATGGAATCGGATTAGGATGCAATTTTTCGACGACAGAAACGCTAG 900
DB 841	ACAGATAGAGAAAGATGGAATCGGATTAGGATGCAATTTTTCGACGACAGAAACGCTAG 900
QY 901	TGGAGCTCTAGAAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTGAGCTTAAAGC 960
DB 901	TGGAGCTCTAGAAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTGAGCTTAAAGC 960
QY 961	AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCC 1020
DB 961	AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCC 1020
QY 1021	AATTCTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
DB 1021	AATTCTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
QY 1081	ACCTGCAGATGGTCTGATGTTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1140
DB 1081	ACCTGCAGATGGTCTGATGTTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1140
QY 1141	AGGAAGTAGTATGTTGATGATTTCCATGCTGATGTTTCCATGCTGATGATGCTGAAAAATGAGAC 1200
DB 1141	AGGAAGTAGTATGTTGATGATTTCCATGCTGATGTTTCCATGCTGATGATGCTGAAAAATGAGAC 1200
QY 1201	CGTCTCCATTTTGTGCTGGGTTTTCGTCAGATGATTCACATGTTTCAATACGGAATAATCC 1260
DB 1201	CGTCTCCATTTTGTGCTGGGTTTTCGTCAGATGATTCACATGTTTCAATACGGAATAATCC 1260
QY 1261	TGATTCTCAAGCTGCCAACAGGAGCTTCGACGACCAAGCTAGACGCAAGGCGCTGG 1320
DB 1261	TGATTCTCAAGCTGCCAACAGGAGCTTCGACGACCAAGCTAGACGCAAGGCGCTGG 1320
QY 1321	AGATGACAGTGTGTCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTTAGG 1380
DB 1321	AGATGACAGTGTGTCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTTAGG 1380
QY 1381	TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGGCTTCTGCTGCTGT 1440
DB 1381	TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGGCTTCTGCTGCTGT 1440
QY 1441	TGTGAGCGCAGAGTA 1456
DB 1441	TGTGAGCGCAGAGTA 1456
RESULT 2	
AX045137	
LOCUS	
Sequence 7 from Patent WO0066739.	
ACCESSION	
AX045137	
VERSION	
AX045137.1	
KEYWORDS	
SOURCE	
ORGANISM	
Chlamydomophila pneumoniae	
Chlamydomophila pneumoniae	
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.	
REFERENCE	
1	
Murdin, A.D., Oomen, R.P., Wang, J. and Dunn, P.	
i(chlamydia) antigens and corresponding dna fragments and uses	
thereof	
Patent: WO 0066739-A 7 09-NOV-2000;	
Aventis Pasteur Limited (CA)	
JOURNAL	
Location/Qualifiers	
FEATURES	
2238 bp	
DNA	
linear	
PAT 24-NOV-2000	

source	1. .2238	
	/organism="Chlamydomophila pneumoniae"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:83558"	
CDS	766. .2238	
	/note="unassigned protein product"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="C1134373"	
	/db_xref="GI:1134373"	
	/db_xref="RENTREMBL:CAC17330"	
	/translation="MNPITGPGPIDERTPPADLSAOGLESAANKSABEAOIRIAGAE APEKESKTSVERMSILRSANVNLMSLADKLIASNSSSRSRSDVDEETATATPT PPTFDYKIQAGTADITFTSLADIQAALVSDAVNTIKDQATDEETAIARAEW ETKNADAVKVGQITELAKYVSDNQLDLSGLKLTDFDLQALQSVASGNKAAELL KEMQNPVPGKTPAQLVSDQDTATQIEKDNARIDRPAFGVAGVANKAKSN NSTISNTDSKAATATATQIAEAKQKFPDPSILQAEQOMVIOAEKDLKNIKPADGSDV PNPPTVGSKQGSISIRVSMLLDAENETASILMSGFRQMIHMENTENPDQAA QOELAAQAAKAAGDSDAALADAKALEALKGAGQOQGLNLALGQIASAAVVA GVLPLOQVWIRARYQAYVEQKLISEDLNSAVDHHHHH"	
ORIGIN		
Query Match	100.0%; Score 1456; DB 6; Length 2238;	
Best Local Similarity	100.0%; Pred. No. 2.3e-301;	
Matches 1456; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATAAATCTTTAAACACAGGCTCGCATTAATTATTAGTGAAGCTTTTTTTTATTTTTT	60
DB	666 ATAAATCTTTAAACACAGGCTCGCATTAATTATTAGTGAAGCTTTTTTTTATTTTTT	725
QY	61 ATAATAAACTAAAGATTTTATTTTATTTTTCAGTTTATTTATGTTAACTATTGCTC	120
DB	726 ATAAATAAACTAAAGATTTTATTTTATTTTTCAGTTTATTTATGTTAACTATTGCTC	785
QY	121 AGTCTCTATAGACGAACAGAGCGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA	180
DB	786 AGTCTCTATAGACGAACAGAGCGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA	845
QY	181 GCGAGTGCAGCAATAAGATGCGGAAGCTCAAGAAATAGCAGGTCGGAAGCTAAGCC	240
DB	846 GCGAGTGCAGCAATAAGATGCGGAAGCTCAAGAAATAGCAGGTCGGAAGCTAAGCC	905
QY	241 TAAAGATCTTAAGCCGATTTCTGAGAGGATGAGCATCTTCTGCTTCTGAGTGAATGC	300
DB	906 TAAAGATCTTAAGCCGATTTCTGAGAGGATGAGCATCTTCTGCTTCTGAGTGAATGC	965
QY	301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG	360
DB	966 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG	1025
QY	361 CAGATCTGAGAGTGGACTCAAGCAGCAGCGCCGACCTACGCTCTCCACCCAGCTT	420
DB	1026 CAGATCTGAGAGTGGACTCAAGCAGCAGCGCCGACCTACGCTCTCCACCCAGCTT	1085
QY	421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTTACCTTCAACATCACT	480
DB	1086 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTTACCTTCAACATCACT	1145
QY	481 AGCTGACATACAGGCTGCTTTGGTGGAGCTTCCAGGATGCTGCTCACTAATATAAGGATAC	540
DB	1146 AGCTGACATACAGGCTGCTTTGGTGGAGCTTCCAGGATGCTGCTCACTAATATAAGGATAC	1205
QY	541 AGCGCTACTGATGAGGAACCGCAATCGTGGGAGTGGGAACCTAAGATGCCATGC	600
DB	1206 AGCGCTACTGATGAGGAACCGCAATCGTGGGAGTGGGAACCTAAGATGCCATGC	1265
QY	601 AGTTAAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCAT	660
DB	1266 AGTTAAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCAT	1325
QY	661 TCTTGACTCTTTAGGTAAACTGACTTCTTCCCTTCGACCTCTTACAGGCTGCTTCTTCCAAATC	720
DB	1326 TCTTGACTCTTTAGGTAAACTGACTTCTTCCCTTCGACCTCTTACAGGCTGCTTCTTCCAAATC	1385
QY	721 TGTAGCAAAATAAACAAGCAGCTGAGCTTTCTTAAAGAGATGCAAGATAAACCAGTAGT	780
DB	1386 TGTAGCAAAATAAACAAGCAGCTGAGCTTTCTTAAAGAGATGCAAGATAAACCAGTAGT	1445
QY	781 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC	840
DB	1446 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC	1505
QY	841 ACAGATAGAGAAAGATGGAATTCGATAGGGATGCATATTTTGACAGGACAGAAACGCTAG	900
DB	1506 ACAGATAGAGAAAGATGGAATTCGATAGGGATGCATATTTTGACAGGACAGAAACGCTAG	1565
QY	901 TGGAGCTGTAGAAAATGCTAAATCTTAATAACAGTATAAGCAACATAGATTCAGCTAAAGC	960
DB	1566 TGGAGCTGTAGAAAATGCTAAATCTTAATAACAGTATAAGCAACATAGATTCAGCTAAAGC	1625
QY	961 AGCAATCGCTACTGCTTAAGACACAAATAGCTCAAGCTCAGAAAAGTTCCCGGACTCTCC	1020
DB	1626 AGCAATCGCTACTGCTTAAGACACAAATAGCTCAAGCTCAGAAAAGTTCCCGGACTCTCC	1685
QY	1021 AATTCTTCAAGAGCGGAACAAATGGTATATACAGGCTGAGAAAGATCTTAAAAATATCAA	1080
DB	1686 AATTCTTCAAGAGCGGAACAAATGGTATATACAGGCTGAGAAAGATCTTAAAAATATCAA	1745
QY	1081 ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACACTAGTTGGAGGCTCCAAAGCAACA	1140
DB	1746 ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACACTAGTTGGAGGCTCCAAAGCAACA	1805
QY	1141 AGGAAGTACTATTGGTAGTATTGCTTTCCATGCTGTTAGATGATGCTGAAAATCAGAC	1200
DB	1806 AGGAAGTACTATTGGTAGTATTGCTTTCCATGCTGTTAGATGATGCTGAAAATCAGAC	1865
QY	1201 CGCTTCCATTTTATGTTGCTGGGTTTCTGTCAGATGATTCATGTTCAATACGGAATAATCC	1260
DB	1866 CGCTTCCATTTTATGTTGCTGGGTTTCTGTCAGATGATTCATGTTCAATACGGAATAATCC	1925
QY	1261 TGATTTCTCAAGTGGCCACACAGAGCTCGCAGACACAGCTAGCAGCAAGGCGGCTCG	1320
DB	1926 TGATTTCTCAAGTGGCCACACAGAGCTCGCAGACACAGCTAGCAGCAAGGCGGCTCG	1985
QY	1321 AGATGACAGTGTGCTGTCAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG	1380
DB	1986 AGATGACAGTGTGCTGTCAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG	2045
QY	1381 TAAAGCTGGGCAACACAGGGCATCTCAATCTTTAGACACAGATGCTTCTGCTGCTCT	1440
DB	2046 TAAAGCTGGGCAACACAGGGCATCTCAATCTTTAGACACAGATGCTTCTGCTGCTCT	2105
QY	1441 TGTGAGCGCGCAGGATA 1456	
DB	2106 TGTGAGCGCGCAGGATA 2121	
RESULT 3		
AX045131		
LOCUS	AX045131	2156 bp DNA linear PAT 24-NOV-2000
DEFINITION	Sequence 1 from Patent WO0066739.	
ACCESSION	AX045131	
VERSION	AX045131.1	GI:11343730
KEYWORDS		
SOURCE		
ORGANISM	Chlamydomophila pneumoniae	
REFERENCE	Chlamydomophila pneumoniae	
AUTHORS	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.	
TITLE	1	
JOURNAL	Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.	
FEATURES	i (chlamydia) antigens and corresponding dna fragments and uses	
source	thereof	
	Patent: WO 0066739-A 1 09-NOV-2000;	
	Avantis Pasteur Limited (CA)	
	Location/Qualifiers	
	1. .2156	
	/organism="Chlamydomophila pneumoniae"	

Handwritten signature and notes:

2007.10.10

Handwritten text: "Chlamydomophila pneumoniae"

```
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
101..2056
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC17329.1"
/db_xref="GI:11343731"
/db_xref="REMTREMBL:CAC17329"
/translation="MVNPIGPGPIDETERTPPADLSAOGLEASANKSAEQRLAGAE
AKPKESKDSVERNSILSRSAVNLADKLGIIASSNSSSTSRSDVSTATAPTP
PPPPDDYKTCQAYDPIFTSLADIQALVSLQDAYVNIKDTAATDEETAIAAEW
ETKNADAVVKGQITELAYSDNOALISLKSFTLLQALLOSVAANNKKAELL
KEMODNPVPGKTPALQSLVDOTDQATQIEKGNFARDYFAGQNSAGVNAKSN
NISINDSAKAAIAFAKQIAEAKQKFPDPILOAEQWVIOAEKOLNIXPADGSDV
PNPITVGGSKQOQSSIGSIRVSMILLDAENETASILMSGFRQIMHMENTNPDSQA
QOELAAQAAKAAAGDGAALADAQKALEAALGKAGQOQILNALGQIASAAVSA
GVPPAAASSIGSVKQLVYKTSKSGSDYKTOISAGVDAYKSIDNAYGHRNDYRDVI
NRVSTPALTRSPVPRARTARGPEKTDQALARVVISGNSRTLGDIVYQVSALOSVQIIQ
SNPQANNEEIRQKQTSVTKPPFGYPPYVQLSNDSTQKFIKLSLESLFAEGSRTAABIK
ALSFETNSLFTQQVLVNIQSLYSYLQ"
```

ORIGIN

```
Query Match
Best Local Similarity 99.9%; Score 1455; DB 6; Length 2156;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAAATCTTTAAAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTT 60
DB 1 ATAAAATCTTTAAAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTT 60

QY 61 ATAATAAACTTAAAGATTTTTTATTTTTTGGATTTTATGTTTATCTTATTTGGTCC 120
DB 61 ATAATAAACTTAAAGATTTTTTATTTTTTGGATTTTATGTTTATCTTATTTGGTCC 120

QY 121 AGGTCCTATAGACGAAACAGACGACACCTCCCGCAGATCTTTCGTCTCAGGATTTGA 180
DB 121 AGGTCCTATAGACGAAACAGACGACACCTCCCGCAGATCTTTCGTCTCAGGATTTGA 180

QY 181 GCGAGTGCGAGCAATAAGAGTGGGAGCTCAAAAGAAATAGCAGTGGGAAAGCTAAGCC 240
DB 181 GCGAGTGCGAGCAATAAGAGTGGGAGCTCAAAAGAAATAGCAGTGGGAAAGCTAAGCC 240

QY 241 TAAAGANTCTAGACCGGATCTGTAGAGCGATGGAGCATCTTCGTTCTGCAGTGAATGC 300
DB 241 TAAAGANTCTAGACCGGATCTGTAGAGCGATGGAGCATCTTCGTTCTGCAGTGAATGC 300

QY 301 TCTCATGAGTCTGCGAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG 360
DB 301 TCTCATGAGTCTGCGAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG 360

QY 361 CAGATCTGCGAGCTGGATCTAACGACGCGACCGGACCTAGCCCTCTCCACCCAGCTT 420
DB 361 CAGATCTGCGAGCTGGATCTAACGACGCGACCGGACCTAGCCCTCTCCACCCAGCTT 420

QY 421 TCGATGATTATAGACTCAAGCGCAACACGCTTACGATATCTTTTACCTCAACATCACT 480
DB 421 TCGATGATTATAGACTCAAGCGCAACACGCTTACGATATCTTTTACCTCAACATCACT 480

QY 481 AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAATATAAAGGATAC 540
DB 481 AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAATATAAAGGATAC 540

QY 541 ACGGGCTACTGATAGGAAACCGCAATCTCGTCGGAGTGGGAAATCAAGATGCGGATGC 600
DB 541 ACGGGCTACTGATAGGAAACCGCAATCTCGTCGGAGTGGGAAATCAAGATGCGGATGC 600

QY 601 AGTTAAAGTTGGCGCAATACAGAAATAGCAAAATATGCTTCGGATAACCAAGCGGAT 660
DB 601 AGTTAAAGTTGGCGCAATACAGAAATAGCAAAATATGCTTCGGATAACCAAGCGGAT 660

QY 661 TCTTGACTCTTTAGGTAAACTGACTTCTTCTGACCTCTTACAGGCTGCTTCTTCTCAATC 720
DB 661 TCTTGACTCTTTAGGTAAACTGACTTCTTCTGACCTCTTACAGGCTGCTTCTTCTCAATC 720
```

```
Db 661 TCTTGACTCTTTAGGTAAACTGACTTCTTCTGACCTCTTACAGGCTGCTTCTTCTCAATC 720
QY 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGT 780
QY 781 CCCAGGGAAACCGCTCGAATTGCTCAATCTTTAGTTGATGATGATGATGATGATGATGATGAT 840
Db 781 CCCAGGGAAACCGCTCGAATTGCTCAATCTTTAGTTGATGATGATGATGATGATGATGATGAT 840
QY 841 ACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATATTTTCAGGACAGACAGCGCTAG 900
Db 841 ACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATATTTTCAGGACAGACAGCGCTAG 900
QY 901 TGGAGCTGTAGAAAATCTAAATCTAATAACAGTATAAGCAACATAGATTGATGATGATGATGAT 960
Db 901 TGGAGCTGTAGAAAATCTAAATCTAATAACAGTATAAGCAACATAGATTGATGATGATGATGAT 960
QY 961 AGCAATCGCTACTCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
Db 961 AGCAATCGCTACTCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
QY 1021 AATTCTTCAAGAGCGGAACAATGGTAATACAGGCTGAGAAAAGTCTTAAAAATATCAA 1080
Db 1021 AATTCTTCAAGAGCGGAACAATGGTAATACAGGCTGAGAAAAGTCTTAAAAATATCAA 1080
QY 1081 ACCTGCGAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACA 1140
Db 1081 ACCTGCGAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACA 1140
QY 1141 AGGAAGTAGTATTGGTAGTATTCGTTGTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
Db 1141 AGGAAGTAGTATTGGTAGTATTCGTTGTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
QY 1201 CGCTTCCATTTTGTGTTGCTGGGTTTCGTTCAGATGATTCACATGTTCAATACGAAAATCC 1260
Db 1201 CGCTTCCATTTTGTGTTGCTGGGTTTCGTTCAGATGATTCACATGTTCAATACGAAAATCC 1260
QY 1261 TGATTCTCAAGCTCCCAACAGGAGCTCGAGACAGCTAGACAGGAGGAAAGCGCTGG 1320
Db 1261 TGATTCTCAAGCTCCCAACAGGAGCTCGAGACAGCTAGACAGGAGGAAAGCGCTGG 1320
QY 1321 AGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGTCTTAGAAGCGGCTTAGG 1380
Db 1321 AGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGTCTTAGAAGCGGCTTAGG 1380
QY 1381 TAAAGCTGGGCAACCAACAGGCGCATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Db 1381 TAAAGCTGGGCAACCAACAGGCGCATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGT 1440
QY 1441 TGTGAGCGCAGGAGT 1455
Db 1441 TGTGAGCGCAGGAGT 1455

RESULT 4
AE001654/c
LOCUS AE001654 10421 bp DNA linear BCT 01-DEC-2000
DEFINITION Chlamydia pneumoniae section 70 of 103 of the complete genome.
ACCESSION AE001654 AE001363
VERSION AE001654.1 GI:4377031
KEYWORDS
SOURCE Chlamydophila pneumoniae CWL029
ORGANISM Chlamydophila pneumoniae CWL029
REFERENCE 1 (bases 1 to 10421)
AUTHORS Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
TITLE Kaiman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R. W.,
JOURNAL Ollinger, L., Greenwood, J., Davis, R. W., and Stephens, R. S.
MEDLINE Comparative genomes of Chlamydia pneumoniae and C. trachomatis
PUBMED Nat. genet. 21 (4), 385-389 (1999)
REFERENCE 2 (bases 1 to 10421)
```

AUTHORS Kalman,S., Mitchell,W., Marathe,R., Lammel,C., Fan,J., Olinger,L.,
Grimwood,J., Davis,R.W. and Stephens,R.S.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1998) Program in Infectious Diseases, University
of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA
FEATURES Location/Qualifiers
source 1..10421
/organism="Chlamydoiphila pneumoniae CWL029"
/mol_type="genomic DNA"
/strain="CWL029"
/db_xref="taxon:115713"
133..2751
/gene="CPn0727"
133..2751
/gene="CPn0727"
/codon_start=1
/transl_table=1
/product="CT619 hypothetical protein"
/protein_id="AAD18866.1"
/db_xref="GI:4377032"
/translation="MSTFSTQRLRTTSGESTRIIKLDHYSGDPDRSPVAINLEELN
SGIYALRLNALQSENVAALNPNTTPTTSTWTDYKHSRPPQASPRAPSQTPT
DIVSAALAVLVLDGLAELVASVTEIDIGALSTISTVQLMASYGLTLTAQSK
VFSVVVPSEKLEHVKQKAAEIOAKOEIKAVLEAKGVSTEEIALIKVPTDVI
AADPFKEFIPELHTYRAKGAPIOENNEAIOLPTPEALPTDNVNEVNGMTLSI
LOALDDAKAPALGCGQEIITLITVLVLDKTTFTKAEFDLITATQIPNTASLKL
YLDRQIAERKITYQNSIQNLSETKRVNENRSMLETQLSMFOQACNCFVTHIS
QANALNATNKYISAVLTSMYEGLLCLSYMYERLADDERAIFDKSYNEVLIPI
VGGSWNGIMAKMAAYQELAEYSLGTAFTSQDQIKAYLQTRGNEFKATRHFNIGD
QMYQFANETVGNCLTTANGAIQPLDGLFTREAMTVGTVEADVNAQALINEFNTA
ATAVLQLOLQALQKADLDLPGKASTENKFAVAMITSESLGALISMLTNSQ
LPKQELKPLIEINFNLANALNSLIQITREFSTSVYSLSVVQSKQGNLF
AGDYELLAAERETIYRDTRKQALINVLGLOKINSLPGATSAQKQEMNAT
YYQSVLNTQNLVLESLAGLQMLTQTSNNKYDKSVFKIESFDWIPTLALESF
LTSQFNISATGLGPIFTQVSDQQTYSQGTQQLNLQNLQNTTIQCEWTLVSTMSQ
VLNGILSQLAGIYSN"
complement (3078..5033)
/gene="CPn0728"
complement (3078..5033)
/gene="CPn0728"
/codon_start=1
/transl_table=1
/product="CHUPN 76 kDa Homolog 1 (CT622)"
/protein_id="AAD18867.1"
/db_xref="GI:4377033"
/translation="MWNPIGPGIDETERTPPDLQAQLEASAAKNSAKAQAQIAGAE
AKPKESKTSVERMSILRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTP
PPTFDYDQTAQAYDTITFTSLADIQAALVSLQDAVNIKDITATDEBTAIAEWM
ETKNADAKVGAQITELAKVADNQAILDQKGLTSPDLLQAALQSVANNKAAELL
KEMQDNPVPGKPAIAQSLVDQPDATQIEKDGNARIDAYFAGONASGAVENAKSN
NSISNIDAKAATATQIAEAKKFPDPILOAEQOMVIOAKEDLNKIKPADGSDV
PNFTTVGGSKQSSIGSRVSMILDAENETASILMSFRQMIHNFENTENPDQAA
QQLAQAARAAGDSDAAALADAKALEALGKAGQCGQLNALGQIASAAVSA
GVPPAAASSTGSSVKQIKYKTSKTSYDQTKQISAGYDAYKSIDAYGRANDATDVI
NNVSTPALTRSPVPRTEARPEKTDQALARVTSNRTGLDGVYSQVSAQSVQMIQ
SNPOANNEETROKITSATVTPPOPGYVQVLSNDSTQKFIKULESLFAEGSRFAAEIK
ALSFETNSLFIQQLVNLGSLYSGLQ"
complement (5145..6413)
/gene="CPn0729"
complement (5145..6413)
/gene="CPn0729"
/codon_start=1
/transl_table=1
/product="CHUPN 76 kDa Homolog 2 (CT623)"
/protein_id="AAD18868.1"
/db_xref="GI:4377034"
/translation="MKQVYQWLASVLLATISGYAELPLSEOKVKSHYVTLIDEVK
DYLKRGKGFTRKQDGLVRIAGDVRAEWLYFEREDIKNPDKDKYKPLPNVRYSEFVL
YIDYABRNWLSKKNWTATAGCENTAAGVDINRAFLGYRKYKNPETRTDFPFEIGRS
GDLDFSEVQFQSGDLGLHIYWTRELSKDYPIQYIVHGGFFVPMNTKHYAVWVEGI
LNRLPKQFVFKCSVVDNNTFVPSSTTEKAATNAMKYCYVQWQLMKHYSQVFWING
QKPELYLYGAPLNAKTTLNGENLAWFTGCTGLGLRKGADWSATVRYEYVA
LSVPEIDVSGIRGNLLKFWFAQIAAANYDPKEANGTNYKGFSAIYVYGIITDLSLSPR

AYGAYSKPANDKLGSDFTFRKFDLGIISAF"
complement (6516..8159)
/gene="mvin"
/notes="synonym: CPn0730"
complement (6516..8159)
/gene="mvin"
/codon_start=1
/transl_table=1
/product="Integral Membrane Protein"
/protein_id="AAD18869.1"
/db_xref="GI:4377035"
/translation="MSRKONEVSLABSIFENILSGTSCSRITGIFERRIAMATYFQADPI
VAAPWLGRTVFFLRKILGGLILEQAFIHFEFLRAQSLDRAAFERFRLIKGSTI
IFTLLIEAVLWVVLQVEGYDMILLTLLPCGIFLMYVNGALLHCENKFPFVG
LAPVYVNIWIFFVIAARSDPRRIIGLSVALVIGFFEWLITVGVWKFLEAKSP
PQHSVRALLAPLSIGLITSSIFQNLASDI CLARYVHEIGELYLMYSLIKYQLPFIH
LFGGVFTVLLPAISRQVCDREHGERGLMKPVLTTMSVMIIMTAGLLILALPGVRV
LYEHGFLPOSANVAIVRVLRGCGASII PMALAIVLSVLVFAQCYVAPVFIIGTALA
NIVLSIVLGRWLKDVSGISYATISITAWQLYFLWYSSKRLPMYSKLLWESRRSIK
VMGTTMLACMITLGLNLITQTYVIFLNPITPLAWPLSSITAQAIAPLSESCIFLAPL
YGFAKLLRVEDLILNASFEYWRGQGLQQRHYMQDTQN"
8143..8319
/gene="CPn0731"
8143..8319
/gene="CPn0731"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAD18870.1"
/db_xref="GI:4377036"
/translation="MSFLIKKPPPLKQDYLFDISPYTSSITIGGSYFKLNKASL
QSTLRLRSISILIS"
8654..9535
/gene="nfo"
/notes="synonym: CPn0732"
8654..9535
/gene="nfo"
/codon_start=1
/transl_table=1
/product="Endonuclease IV"
/protein_id="AAD18871.1"
/db_xref="GI:4377037"
/translation="MKVLPPSPITLLGAHTSTAGLKNAIYEGRDIGASTVQIFLTAQ
RQWRALKEVIEDFAALKEITDLSYIMSHAGYLINPGADPVPVILEKSRIGIYQELL
DCITLGSFVNFHPHGAALSKESDKWKIVSPSQSAPLDPSPVILVLETTAGGT
LIGNSFEELGYLVQNLKNOICGVCDVTCHIFAAGYDITSPQGWEDVLEDFYVGLS
YLAFHLNDSMFPLGANKORHAPLGRGYIGKESFKELMTDTRKIPKYLETPGGEN
WQKEIGELLKFSKNRDS"
complement (9654..10283)
/gene="rsa"
/notes="synonym: CPn0733"
complement (9654..10283)
/gene="rsa"
/codon_start=1
/transl_table=1
/product="S4 Ribosomal Protein"
/protein_id="AAD18872.1"
/db_xref="GI:4377038"
/translation="MARYCGPKNRVARRFCANIFGRSRNPLKKKPHPPGQGMQRKKK
SDYGLQLEBKQKLKACYGMIMKQVKAKEVTHKQNVQAMFELFERECRLDNVYRM
GPAKTIFAAQQLVAHGHIILVNGRRVDRRSFELRPGMQISLKEKSKLQSVKDALESKD
ESSLSYISLDDKTFKGLLVSPQDQIEAQLPLPINISVVCFLSHRI"

ORIGIN

Query Match 99.9%; Score 1455; DB 1; Length 10421;
Best Local Similarity 100.0%; Pred. No. 3.2e-301;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATAAATCTTTAAAAACAGCTCGCATTAATTAATTAATAGTAGAGACTTTTATTTATTTT 60
Db 5133 ATAAATCTTTAAAAACAGCTCGCATTAATTAATTAATAGTAGAGACTTTTATTTATTTT 5074
Qy 61 ATAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120


```

DRSHYTPERANILWHPDAFDNDLLISDIKRLKNNNEIQAQVDFVLGNRSKTEJET
IYPSKVLIVFENQELRLMDIRIFVDYTDADERILRMVDRVQEQDSVDCIMS
RYSLSMVKMEKEIEPTKVDYIIVHGNYRQNVVNTLSQIKKNHLENALESDETVM
VNSK"
/complete(1047. .2021)
/gene="CP0012"
CDS
/complete(1047. .2021)
/gene="CP0012"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37908.1"
/db_xref="GI:7188951"
/transl_table=11
/translation="MERKYYALAYYYITRVDPHEEIALHKLFLEDLDVSCRIVYBQ
GINDQSGYGPHEALYMWQLKERNFSKIKFKIHKENTPPRITVKYKELAAIGE
VDLSQAKHTSPQEWHEKLOENRCLLLDVNNVWEMKIGHEDNATLPDIOTFRFPPEYA
EKLAQECDDPTTVMYCTGIGRCELSPVLEKGFKEVYQLDGGVIAIYQQVGTGKW
LGLFVDDRLAIPIDESDPVAPAECCHCQTPSDAYNCANTDCNALFLCCDCBCH
QHQCCGECQSQFVRKFDPSRGKPFRAHLCEISENSESACCLII"
2238. .2867
/gene="CP0013"
CDS
/complete(2867. .2867)
/gene="CP0013"
/notes="similar to SP:P81289; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein S4"
/protein_id="AAF37909.1"
/db_xref="GI:7188952"
/translation="MARYCYGPKNRVARRFGANIFGRSRNPLLLKPPHPPGQHGQRKK
SDYGLQEEQKQKACQYKIMEXQLVKAQFVHKQGNVQAFMEFPERCECLDNNVYRM
GFAKTIFAAQLVAHGHLVNGRVRDRSPFLRPGMOISLKEXSKRLQSVKDALESKD
ESSLPYSYISDKTGFKCELLVSPEDQIEAQLPIPINISVVCFLSHRT"
/complete(2986. .3867)
/gene="CP0014"
CDS
/complete(2986. .3867)
/gene="CP0014"
/notes="similar to SP:P12638 GB:M22591 PID:146954
PID:405898 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="endonuclease IV"
/protein_id="AAF37910.1"
/db_xref="GI:7188953"
/translation="MKVLPPLPSIPLGATSTAGLKNATVEGIDICASTVQIFATNQ
RQQRALKEVEDFKAALKETDLSYMSHAGYLINPAPDPVILEKSRIGIYQEL
DCITLGLSFVNFHPGALKSKEDCMNKIVSSFSQAPLFDSPPLVLLLETTAGQT
LIGNFEEGLYVNLKNQIPIGVCDTCHIFAGYDITSPQWEDVINEFDYVGLS
YLRAFHLNDSNFFLGNKNDRHAPLEGYIGKESFKFLMDTERTRKIPKYLETPGPN
WQXIEGLLKFKNRDS"
/complete(4202. .4378)
/gene="CP0015"
CDS
/complete(4202. .4378)
/gene="CP0015"
/notes="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF37911.1"
/db_xref="GI:7188954"
/translation="MSFLIKRPSPPPLKQDYLFDISPYTSSEITIGGSYFLKNKASL
QSSTLRISISIIIS"
4362. .6005
/gene="CP0016"
CDS
/complete(6005. .6005)
/gene="CP0016"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"

```

```

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37912.1"
/db_xref="GI:7188955"
/translation="MSKKNESVLARSIFNLSGTCSRTIGFIREIMATYFGADPI
VAAFGLGRTVFFRLKILGLLEQAFIPHEFLRAQSDLRDAFFFRFRLKGGTI
ITFLTEAVLWVLOVEEGTYMILLTLLPCGIFLMMYNNVALLHCENKFPVGG
LAPVVNIIWIFVIAARSDPRRIIGLSVALVIGFFEWLITVPGVWKFLFEAKSP
POEHSRALLAPLSGILTSIFOLNLLSDICLARYVHEIGLYPLYSKLYQLPPIH
LFGGVTLLPAISRCVOREHGLKMKFVLTITMSMIITWAGLLILLALPGVRV
LYEHGLFPQSAVYAVLVRGASIIIPALAPLVSVLPYQAQYAVPLFICGTALA
NIVLSLVGRWLKNDVSGISYATISAWQVLPFLWYSSKRLPMYKLLWSEIRRSIK
NVGTMACMITGLNKLTQTVTTFINFLTPLAWFLSRTITQAIAIFSESCIFLAF
EFGAKLLRVEDLINLASFYWRGQRLQORQHVMDTON"
6108. .7376
/gene="CP0017"
CDS
/complete(7376. .7376)
/gene="CP0017"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37913.1"
/db_xref="GI:7188956"
/translation="MKQVYQMLASVLLALITISYAEPLSEQKYSHTYTLDEVK
DYLKRGFVETRKQDGLVRIAGDVRARWLYFREDIKNPSDKKYNPLPNRYSEFVL
YDYRAERNWLSKNNMTAIGCENTAAGVDINRAFLGYRFPKNPRTDFFNVEIGRS
GLGDLPESSVQSQSNFDGLHYWTRLSKDYQVYLHVGSPVNNTKHYAVVBSGI
LARKLPQFVKSVVDWNTFVSESTTEKATNANKYKVCWQWLGVKHSQVPWING
QKRLPLYLGAFLMNLAKATKTLNKENLWPIGGTGLGRKAGWSATFVRYVEA
LSVPEIDVSGIGRGNLKFQFAQIAAYDPKEANGFTNYKGFSAIYMTGITDLSFR
AYGAYSKPANDLGSDFTRKFDLGIISAF"
7488. .9443
/gene="CP0018"
CDS
/complete(9443. .9443)
/gene="CP0018"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37914.1"
/db_xref="GI:7188957"
/translation="MVNP:GPGPIDETERTPPADLSAQGLEASAAKSAEQAQIAGAE
AKPKESKTSVERWSILRAVNAVMSLADKLGIASSNSSSTSRSDVSTTATPTP
PPPTDYKTQACTAYDTIFTSLADIAALVLODAVNTIKDTRATDEETAIAEAW
ETKNADAVYGAQITELAKYASDNQAILSLGKLTSLLOALLOSANNKAAELL
KMQDNPVYFGKTPALAQSLVDQTDATATQIEKGNARDAIFAGQASGAVENAKSN
NLSNIDSAAKALATARTQIAEAQKFPDPSILOEAQWVIQAEKDLKIKPADSDV
PNPPTVGSGKQSGSIGIRYVSMLLDAENETASITMSGFQIMHFNTEPDSQAA
QBELAAQAAKAAAGDDSAALADAKALEALGKAGQQGLNALGQITASAAYVSA
GVPPAAASIGSGSKOLYKTSKSDYKTOISAGDAYKSINDAYGRANDATROVI
NVSTPEALTESVPRATEARGPEKTDQALRVISGNSRTLGDVYSOVSALQSWQILIQ
SNPQANNEIRKLTSAVTKPQFGPYVQNDSTQKFIKLESFASGSRFAEIK
AUSFETNSLIFIOQVLVNLISLYSGYLQ"
/complete(9770. .12388)
/gene="CP0019"
CDS
/complete(9770. .12388)
/gene="CP0019"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37915.1"
/db_xref="GI:7188958"
/translation="MSTFSIQNRLRTISGESTRIIKLDHKYSGFDRSPVAINLEEN
SGIYALRHLNALQENTNVAAALNPNNTIPTFTWTDYKHSRPOASSPRAPSSQPTT
DIVSAALALVLDGGLAEVAVTEIDGLASTISTIRQLMASYGLTTLTAEOEK
VWFSYSPSEKNLLEHVQKKAARTQAKERIKAVLEAKGVSTEEIALLKEYPDYI
ADDFKEFIEEPLHYTRAKVAPIQEMENALQLLPTFPALIPDNVNVNGNTLSTI

```

LQAIIDAIKQAPALGGBQEIITLQTLVPLVDKTFITKAEPDLLYATQLPNTASUKI
 LQTLDAEYRGKTIYQKQNSQTSKTRVKNNSMLQTSYVWVQOACQNSWIS
 QANALANTIKYISAVLTISWEMVGGLLCSYMYRLLADDEKALPKS VNEYLPTHI
 VUGSWNGHIAKQAKQELAEYS:GTAVTSQDLKAYLQTRGNEFKATRPHFNIGD
 QMYQFANETVFGNLTANGATQPLDGFITRNFVTEADYSNAQRLINFNFA
 ATAVHLOICLAELKQADDLPCKSGASTENKTRFVAVAWITSISGLGAILMSLNS
 ATAVHLOICLAELKQADDLPCKSGASTENKTRFVAVAWITSISGLGAILMSLNS

Query Match	99.9%	Score 1455	DB 1	Length 12548
-------------	-------	------------	------	--------------

Query Match	99.9%;	Score 1455;	DB 1;
Best Local Similarity	100.0%;	Pred. No. 3.1e-301;	

Best local similarity 100.0%, read: NO. 3-10-2007,
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATATAATCTTTTAA	60
Qy	1	ATATAATCTTTTAA	60
Db	7388	ATATAATCTTTTAA	7447
Qy	61	ATAATAAACTAAA	120
Db	7448	ATAATAAACTAAA	7507
Qy	121	AGGTCCTATAGA	180
Db	7508	AGGTCCTATAGA	7567
Qy	181	GCGCAGTGCAG	240
Db	7568	GCGCAGTGCAG	7627
Qy	241	TAAAGAACTC	300
Db	7628	TAAAGAACTC	7687
Qy	301	TCTCATGAGTCT	360
Db	7688	TCTCATGAGTCT	7747
Qy	361	CAGATCTGCAG	420
Db	7748	CAGATCTGCAG	7807
Qy	421	TGATGATTA	480
Db	7808	TGATGATTA	7867
Qy	481	AGCTGACATAC	540
Db	7868	AGCTGACATAC	7927
Qy	541	AGCGGCTACT	600
Db	7928	AGCGGCTACT	7987
Qy	601	AGTTAAAGTTG	660
Db	7988	AGTTAAAGTTG	8047
Qy	661	TCTTGACTCTT	720
Db	8048	TCTTGACTCTT	8107
Qy	721	TGTTAGCAAA	780
Db	8108	TGTTAGCAAA	8167
Qy	781	CCGAGGAAA	840
Db	8168	CCGAGGAAA	8227
Qy	841	ACAGATAGAGA	900
Db	8228	ACAGATAGAGA	8287
Qy	901	TGGAGCTCTA	960


```

CDS
/locus_tag="CpB0543"
126..797
/notes_tag="CpB0543"
/notes="similar to hypothetical protein CT433 - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98472.1"
/db_xref="GI:33236384"
/translations="MTLYGLNOKTARKYOAHYLPILTPPYAKSTPQNKRALQFLPO
ATHVILTPSGSTHLFTRMSTLSKATLTKTVLCIGESTKERLISFLGVKXVAVTO
EIAEGLPQLQALPSSARILYPHSSLARPVIREFLYNRFTFSYPHYTVKPKLXNNI
LSYKXKILLSPSTVRAPAKLIFRPFKEKTYWCQORMTLQEFQRFSSQKQVSLLETGK
SRUSP"
gene
862..1194
/locus_tag="CpB0544"
862..1194
/locus_tag="CpB0544"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98473.1"
/db_xref="GI:33236385"
/translations="MASSATPGDGTAPSLFPLATPRYNFKLALFTVIALVWIAL
IATIALGICHLPCSFIFLTAFLPISFYICSHYARNVYIALDVVDHSHKLQDMRS
HSPFSDR"
gene
1242..2321
/locus_tag="CpB0545"
1242..2321
/locus_tag="CpB0545"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98474.1"
/db_xref="GI:33236386"
/translations="MSPQSTESSQVSLSVYPRDKELAPKQFTIAKISTLAILASL
ALGAVAGISLTVLQNPVLALILITAPLPSVTVFVYHOMTSKVSQKVLQBNFK
PLGKAVQKNVDCYSNEMQYNHNLAPKFKVALQTDASQFFQFTITGLURVIEKNQST
GIIFNPGVNLINDATNLJUSTILYSLDKDSYMDCKQREGDPKAGEOPFSPTEVVR
VKLPNALDQTNLNLSSAEKKSILFTFLGHVCGPKSEELPNQOEYRQALLAYENCL
KAAIESHAALVALPLFTSVYVPEPEILPKEGTFYWDNQQAFCRALLDAIQNTALR
YQPSLLVILQDPFNTIESQSRSEE"
gene
2661..3425
/locus_tag="CpB0546"
2661..3425
/locus_tag="CpB0546"
/notes="similar to hypothetical protein CT398 - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98475.1"
/db_xref="GI:33236387"
/translations="MHDALLSILAIQELDIKMLRMVKKEHOKELAKVQSLKSDIRR
KVQEKLEMLNATQIRDGNRIQEISEQINKLENQAAVKKMDENALTOEMTANK
ERSLBHQSLDKQAGEDLLVSLKESLASTENSSSVIEKIFFSIKKINEGKAL
LEORTLKHAINPELLSILYERLNNKDRVVFEINRVCVGGCHVILTPQHENLVRRKD
RLIFBHCGRILYQBSQVNAQENSTAKRRRRRAAV"
gene
/genes="kpsF"
/locus_tag="CpB0547"
complement(3846..4835)
/genes="kpsF"
/locus_tag="CpB0547"
complement(3846..4835)
/genes="kpsF"
/locus_tag="CpB0547"
complement(3846..4835)
/notes="similar to kpsF protein - Escherichia coli"
/codon_start=1
/transl_table=11
/product="kpsF"
/protein_id="AAP98476.1"
/db_xref="GI:33236388"
/translations="MPSPMISTDVCODILGKQKEAVDFFQAFQPKEMQLAEKILGH
SGWVFTSGVSGCVARKVATLQSLSERALFFSPVDLLHGDILGLVSPGDIVCLFSKS

```

```

GETQELLDTVPHLKSRRAILVAITSPYNLAALSLLVILPSVABELDFNLIPTNST
TCOMTFDGLAMLLFHSRGSLSITYGKNHPSGOVMKANGKVDKFPFTEVDFCHLG
DKVFSLEVSAYGCGVCIVDPQFRIMGIFTDGLRRLSLASYGGSVLISLEKVMTA
NPRCIFEDSDIAIALQLEMSSSPQAVLPLVDNENRHRVTGLLHMHTLAKAGLL"
complement(4862..6043)
/locus_tag="CpB0548"
complement(4862..6043)
/locus_tag="CpB0548"
/notes="similar to dihydrolipoamide S-acetyltransferase - Bacillus stearothermophilus"
/codon_start=1
/transl_table=11
/product="dihydrolipoamide S-acetyltransferase"
/protein_id="AAP98477.1"
/db_xref="GI:33236389"
/translations="MIFEFPPKIGTSSGSSIVRWLKNLGDHVARDEPLIEVSTDKI
ATELPSPKAGRLVRFVNEGDEVAAGDVLGLIELEISEADDESTCPLTSCSTKSEA
GSSSSVWFSPAVLSLAQREGIQLNLOKTAGTGGKGRVTRODLEAVISESQVSIPE
IPQGEVNRIPMSPLRALIASLSKSSDEVPHASLVVDVDTLMLHLSGERQRFLOTH
GVKLTIFIVQCLAQTLQFPELLSGDGTIIVMKKSVNVAVNLNKEGVVIVH
NQDRGLVSIKALADLSRRLNKLDSEVDQGSVTVNFGMTGALIGMPIIRYPEV
AILGICTIQKRVVVRDDDSLAVRKMVYVTLTFDHRVLDGIYGEFILTSLKNLESVTM
G"
complement(6057..7301)
/locus_tag="CpB0549"
complement(6057..7301)
/locus_tag="CpB0549"
/notes="similar to proton/sodium glutamate symport transporter family"
/codon_start=1
/transl_table=11
/product="proton/sodium glutamate symport transporter"
/protein_id="AAP98478.1"
/db_xref="GI:33236390"
/translations="MKLMKIFILGFVGVTLGLVLEDKAIFFKPIGDIENLLSMVVY
PIVFCSMVLGIASISDMKKGLGVLGTALAVIGLCPAWIFSPGNGCDFA
QAQMSDAVTDIDSNKTAAYFLSIIAQVPSNPVRSFAGNLIQIIFALFLGIARL
SERGPVRFIDGSEIMLRMWNIMSPAPGVGASMAWISGNHGLGVWLQGLKIL
ATYLCATFATLVGLVRFCKMSFKFLSSMDAISCAVSTASSATILPVTMRCVS
KNLGVSAEVSFGVLPGLGATVNMNGTAIFQGMAAVTAQAYNCPCLSSLLLVVATF
SVAGSAGVGGGMITIGLSVASGLP IQGIALIGLDRIDVIGTPMNLGDAVVATY
VASGELSPYESIQESVETT"
complement(7301..8398)
/genes="ycaH"
/locus_tag="CpB0550"
complement(7301..8398)
/genes="ycaH"
/locus_tag="CpB0550"
/notes="similar to ycaH protein - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="YcaH"
/protein_id="AAP98479.1"
/db_xref="GI:33236391"
/translations="MKKRPSLTFLFYRRVTIAISLEGILGWGLSGLLSKVFAFLVA
CNRFSWSPVPEARSTVIVGKIVGAGKCTPVLWLABRLRGVSCGLSRGVKSO
SRQKXLIYVDSKVSASVGBEPLMAKLEPGSVVWHKORRISAAARAEFGILL
DGLQRLKHQKDEIAVAVNQDPLGGRAPFPKGRILRDFPLKTLVTDIIVNGGKEAG
TVGVVSNAPQIFVKPTIASVVMYNGERIPKEALRELVRGVFCGLFPGQGLNMLRE
EGIHILGYLLPDHAAITKKELNYFCQOMAMRQGGQLLCTEKDQSVKLPRLSEVSLIP
IAKVMRLSVNODDTLSLNLNLEQIHKNRN"
complement(8596..9393)
/genes="spoU"
/locus_tag="CpB0551"
complement(8596..9393)
/genes="spoU"
/locus_tag="CpB0551"
/notes="similar to tRNA methylase spoU - Escherichia coli"
/codon_start=1
/transl_table=11
/product="tRNA methylase"
/protein_id="AAP98480.1"
/db_xref="GI:33236392"

```


the United States
J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)
20298986
10839753
2
Shirai, M., Hirakawa, H., Kimoto, M., Takachi, M., Kishi, F., Ouchi, K.,
Shiba, T., Ishii, K., Hattori, M., Kuhara, S. and Nakazawa, T.
Comparison of whole genome sequences of *Chlamydia pneumoniae* J138
from Japan and CWL029 from USA
Nucleic Acids Res. 28 (12), 2311-2314 (2000)
20330349
10871362
3 (bases 1 to 300550)
Shirai, M.
Direct Submission
Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
School of Medicine, Department of Microbiology, 1-1-1
Minamikogushi, Ube, Yamaguchi 755-8505, Japan
(E-mail: mshirai@po.cc.yamaguchi-u.ac.jp, Tel: 81-836-22-2227,
Fax: 81-836-22-2415)
On or before Sep 15, 2000 this sequence version replaced
gi:6172398, gi:6172300, gi:6172396, gi:6172398, gi:8978889.
Location/Qualifiers
1. 300550
/organism="Chlamydia pneumoniae J138"
/mol_type="genomic DNA"
/strain="J138"
/db_xref="taxon:138677"
/note="synonym: Chlamydia pneumoniae (strain J138)"
/complement (21..794)
/gene="dap"
/complement (21..794)
/gene="dap"
/codon_start=1
/transl_table=11
/product="diaminopimelate epimerase"
/protein_id="BAA98725.1"
/db_xref="GI:8978890"
/translations="MAPYSPSTISKYFIYSAGNRFLGTLPEVEDVRLCOETRVD
GFILYLPSSCADAQLIIFNSDGRPMCGNLGALCAHLASQKSDISVSTDSGLYS
GIYFMDRVLDMLADRASVHRLSRPDLPEKVVCIHTGVHAWVILFEISTDL
SILGFLPRLVHTQSPGVNFMVQILGHCOLRVRTYRVEGVEGTAACTGALASLV
SNYSWKSESQIHTWAGELMTVSQNRGRVYLGVSRTDL"
/complement (763..1338)
/gene="clpP 1"
/complement (763..1338)
/gene="clpP 1"
/codon_start=1
/transl_table=11
/product="CLP protease"
/protein_id="BAA98726.1"
/db_xref="GI:8978891"
/translations="MADGEVHKURDIEKLLERARVFPSEPTKSAIDAIKJWYL
ELDKPKPIVFINSGSVSDAGVWDQIKMLTSPVTVTVGLAASMGSVLSLCAAP
GRFAPTHIMQHSIGPTGQATDLDIHAREILKTKARIIDVVEATNQPDII
EKAIIDRMWMTANEAKDFGLDGLIFSNDL"
/complement (1358..2917)
/gene="glyA"
/complement (1358..2917)
/gene="glyA"
/codon_start=1
/transl_table=11
/product="serine hydroxymethyltransferase"
/protein_id="BAA98727.1"
/db_xref="GI:8978892"
/translations="MLKVFKEFKFAIVEIFTKWAVVLSLLKFLFNASKKQGSILAS
TYALADHLNAPPSIGERIIDELQSRLHMIASENYSSVLQAMGNLLTDKYC
EGSPFKRYSCCNVDIAEWECETAKELPAADCAVOPHSAGDANLLAUMALTTHKY
GSPAVSLGYKVTNLTETERYTLKAPMSSCVCLGSLNSGGHLTGNTVLNWSKIM
RCPFDVNPTECFDVAIEIRLAKYKPKVLIAGYSISRRKNFALVKAEDCGSVL
WVDMAHPAGLVAGVDEENPIPIADIVTTTHKLRSGRGNLNLATREYESTLGT
CLPMGGPFLPHVIAAKTVALKEALSVDFFKYAHQVNNRRLAERLISHGLRLTGT
DNHMMVLDLGLSGIKIADILSSVGIAVNRNLSLPSDITGKRWDTSGIRLGTALTL

GMGIDEMERVADIIVKVLNIRLSHVGESSKKNKGELPEATAQAQEARVNRNLLRFP
LYPEIDLEALV"
/gene="CPj0522"
3107..3778
/gene="CPj0522"
3107..3778
/codon_start=1
/transl_table=11
/product="Ct433 hypothetical protein"
/protein_id="BAA98728.1"
/db_xref="GI:8978893"
/translation="MTLYMGLNQTKARKYQAHVLPILTLFPYAKSPONKRALQFLPQ
ATHVILTPSSSTHLFLSMTSLSKATLTKTVLCIGESTKERLSFLGQVKKVATQ
EIAEGIFLLOALPSSARILYEHSSLARVIREFLYNRTFFSYPHYTVKPKLKNIL
LSKIKIIFLTPSTVRAFAKIPFPFKTYICQGRNTLOEQFKFSQKQVSLLETGK
SRTSP"
3843..4175
/gene="CPj0523"
3843..4175
/gene="CPj0523"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98729.1"
/db_xref="GI:8978894"
/translation="MASSATPGDGTAPSLFPAPTEPRYNEKLALFVTLIALVWIAL
IATTAIGLCIHPLCSFIFLTAIPFYISRYICSHYARNVYIALDVDPHSLKQDMRS
HSPIFSDR"
4223..5302
/gene="CPj0524"
4223..5302
/gene="CPj0524"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98730.1"
/db_xref="GI:8978895"
/translation="MSGPSRTSSQVSVLYVPRDEKAIAPKKOFTIAKISTLAILASL
ALGALVAGISLIVLGNPVFLALLITALTALFSVVTLVVHQTSSKVSNNQKVLQNFK
PGLKQWERNKVCYNEMOYNNHLNPKFKVALQIDASQFPFTFLTGURVIEKNSGT
GIIFNPGLNDLNTATNLSTLYLTKDKSVWDTCKQREGPAKGEPPFTVRV
VKLPBALDPTNLNLSAEEKSLPTFLHGVCGPKSELPNQOEYRQALLAYENCL
KAAISHAAIVALPLFTSVYEVPEILPKEGTFYWDNQTAFCRALLDAIQNTALR
YPCRSLLVILQDPFNTIESQSRSE"
5642..6406
/gene="CPj0525"
5642..6406
/gene="CPj0525"
/codon_start=1
/transl_table=11
/product="Ct398 hypothetical protein"
/protein_id="BAA98731.1"
/db_xref="GI:8978896"
/translation="MHDALLSLAIQELDKIMRLMRVKKHQKELAKVOSIKSDIRR
KYQEKLEHMLNKTQIRDCENRIQISLEQINKLQAAQKQKMDENALTOBMTANK
ERRSLEHNSLDLMDQAGCEDLIVLSKESLSSNSSSVIEKEIPSIKINEGKAL
LEORTLKHATNPELLSIVYERLNNKKORVVPIENRVCSCGCHVILTPQHENLVKDD
RLIFCHCSRILYWQESQVNAENSTAKRRRAAV"
/complement (6827..7816)
/gene="kpsF"
/complement (6827..7816)
/gene="kpsF"
/codon_start=1
/transl_table=11
/product="GutQ/kpsF family sugar-P isomerase"
/protein_id="BAA98732.1"
/db_xref="GI:8978897"
/translation="WPSMISTVCQDILGKQKEAVDFEFQAFQKQKAMQLAEKILGH
SGWFPFGVSGCVARKLVAIQLSIRALFFSPVLLHGLDGLVSPGDIIVCLPSKS
TCQMLDITVPLHKKRRALLVAITSMPSYNLAALSDLVVILPSVAELDFNLIPNST
TQMTFGDPLHKLRRGSLSTYGNKPSQGVQMGKANGKVPFKTFVFCFLHG
DKVFSLEFVSAYGGCVIVDPQFLMGIFTDGLRRSLASVYGGVSEVLSLEKYMTA
NPRCITEDSDIAIALQLMESSSPVAVLVLDNEENRHTVGLLHMTLAKAGLL"

gene	complement (7843. .9024) /gene="sucB_2" complement (7843. .9024) /gene="sucB_2" /codon_start=1 /transl_table=11 /product="dihydrolipoamide succinyltransferase" /protein_id="BAA98733.1" /db_xref="GI:8978898" /translation="MIFEFPPKIGETSSGGSIVRLKLNLDHVARDEPLIEVSTDKI ATELPSKAGRLVFCNVEGDEVASGVILGILEEISEADDESTSCPLTSCETPKSA GSSSPSVFSPAVLSAQREGIGLNDLQKTAGTGKGRVTPDLEAVISBQOVLPE IFQEVNRIPKSPRRRLASLSKSDSEVPHASIVDVVDLWNLISGERQRFDLTH GKULITFVIOCLAQLRPFLNGLSDGTTIVMKSVNVGVAVNKNKSGVVPVLIH NCQDGLVSIKALADLSRANLKDLPSEVDGDSVTIVFGMTGALIGMPITRIPEV AILGIGTIQKRVVDDSLAIRKMWYVTLTFDRVLDGIYSEFLTSLKNRLSEVTM G"	Db	217278	TGATGATTATAAGACTCAAGCGCAACAGCTTAGGATACTATTCTTACCTCAACATCACT	217219
CDS		Qy	481	AGCTGCACATACAGCGTCTTTTGGTGGAGCTCCAGGATGCTGTCTCACTAATAATAAGGATAC	540
		Db	217218	AGCTGCACATACAGCGTCTTTTGGTGGAGCTCCAGGATGCTGTCTCACTAATAATAAGGATAC	217159
		Qy	541	AGCGGTCTAGTAGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGATGCCGATGC	600
		Db	217158	AGCGGTCTAGTAGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGATGCCGATGC	217099
		Qy	601	AGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGAT	660
		Db	217098	AGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGAT	217039
		Qy	661	TCTTGAATCTTTTAGGTAAATCTGATCTCTTCGACCTCTTACAGGCTCTCTTCTTCAATC	720
		Db	217038	TCTTGAATCTTTTAGGTAAATCTGATCTCTTCGACCTCTTACAGGCTCTCTTCTTCAATC	216979
gene	complement (9038. .10282) /gene="gltT" complement (9038. .10282) /gene="gltT" /codon_start=1 /transl_table=11 /product="glutamate symport" /protein_id="BAA98734.1" /db_xref="GI:8978899" /translation="MKGMKPIGILFVGVTGLVLEDKAIFPKPIGIDFLNLLSMVVY PLVFCVSLGASISDMKMLGRIKISVGLYLGTALAIIVGLCFWIFSPGNCDEFA QAQMSDAVTIDSNKTAAYFLSIQAQVPSNPVRSFAEGNLIQIIIFAILPLGALRL SGERGPVERIDGFSSEIMLVNMMKMSFAGYGSAMWISGNHGLGLVMOLGKFI AYVLACILPHATLVPGLVRFCKMSFKFLSSMDATISCAVSTSSATLPTVTKRVS KULGVSEVSGFVPLGNTAVNMGTAFQGNAAVFLAQAAYNCPLISLSLLLVTAIF SAVSGVPGGMILGSLVASVGLPIQGIALLAGIDRLRIDVIGPNNILGDVAVVY VASGEGLSPYESIKQESVEIT"	Qy	721	TGTAGCAACAATAACCAAGCGCTGAGCTTCTTAAAGAGATCAAGATAACCCAGTAGT	780
CDS		Db	216978	TGTAGCAACAATAACCAAGCGCTGAGCTTCTTAAAGAGATCAAGATAACCCAGTAGT	216919
		Qy	781	CCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC	840
		Db	216918	CCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC	216859
		Qy	841	ACAGATAGAGAAAGATGGAATGCGATTAGGATGCGATATTTTCAGGACAGAACCGTAG	900
		Db	216858	ACAGATAGAGAAAGATGGAATGCGATTAGGATGCGATATTTTCAGGACAGAACCGTAG	216799
		Qy	901	TGGAGCTGTAGAAAATCTAAATCTAATAACAGTATAAGCAACATAGATTCAAGCTAAAGC	960
		Db	216798	TGGAGCTGTAGAAAATCTAATAACAGTATAAGCAACATAGATTCAAGCTAAAGC	216739
gene	complement (10282. .11379) /gene="ycalH" complement (10282. .11379) /gene="ycalH"	Qy	961	AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC	1020
CDS		Db	216738	AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC	216679
		Qy	1021	AAATCTTCAAGACGGGACCAATGGAATACAGGCTGAGAAAGATCTTAAATATCAAA	1080
		Db	216678	AAATCTTCAAGACGGGACCAATGGAATACAGGCTGAGAAAGATCTTAAATATCAAA	216619
		Qy	1081	ACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACTACAGTTCGAGGCTCCAAAGCAACA	1140
		Db	216618	ACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACTACAGTTCGAGGCTCCAAAGCAACA	216559
		Qy	1141	AGGAAGTAGTATTCGGTAGTATTCGTGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC	1200
		Db	216558	AGGAAGTAGTATTCGGTAGTATTCGTGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC	216499
		Qy	1201	CGCTTCCATTTTGTGATGCTGGGTTTCGTGATGATTCAGATGATTCATACGGAATATCC	1260
		Db	216498	CGCTTCCATTTTGTGATGCTGGGTTTCGTGATGATTCAGATGATTCATACGGAATATCC	216439
		Qy	1261	TGATTTCTCAAGCTCCCAACAGGAGCTCGCAGCACAAGCTAGACGCGAAAGCCGCTGG	1320
		Db	216438	TGATTTCTCAAGCTCCCAACAGGAGCTCGCAGCACAAGCTAGACGCGAAAGCCGCTGG	216379
		Qy	1321	AGATGACAGTCTGCTGAGCGCTGGCAGATGCTTCAGAAAGCTTTAGAACGGCTCTAGG	1380
		Db	216378	AGATGACAGTCTGCTGAGCGCTGGCAGATGCTTCAGAAAGCTTTAGAACGGCTCTAGG	216319
		Qy	1381	TAAAGCTGGGCAACAACAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGT	1440
		Db	216318	TAAAGCTGGGCAACAACAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGT	216259
		Qy	1441	TGTGAGCGCAGGAGT 1455	
		Db	216258	TGTGAGCGCAGGAGT 216244	
RESULT 8					
CHT76KDA					

LOCUS CH776KDA 2148 bp DNA linear BCT 17-MAR-1994
 DEFINITION Chlamydia pneumoniae 76 kDa protein gene, complete cds.
 ACCESSION L23921
 VERSION L23921.1 GI:435961
 KEYWORDS 76 kDa protein.
 SOURCE Chlamydia pneumoniae
 ORGANISM Chlamydia pneumoniae
 Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 1 (bases 1 to 2148)
 Perez Melgosa, M., Kuo, C.C. and Campbell, L.A.
 Isolation and characterization of a gene encoding a Chlamydia
 pneumoniae 76-kilodalton protein containing a species-specific
 epitope. Immun. 62 (3), 880-886 (1994)
 JOURNAL Infect. Immun. 62 (3), 880-886 (1994)
 MEDLINE 94156481
 PUBMED 7509320
 COMMENT Original source text: Chlamydia pneumoniae DNA.
 FEATURES
 Location/Qualifiers
 source 1..2148
 /organism="Chlamydia pneumoniae"
 /mol_type="genomic DNA"
 /db_xref="taxon:83559"
 CDS 1..2148
 /codon_start=1
 /translation="MTKKHVAWVVEGILNLPKOFFVKVCSVDWNTFVPSSTTEKA
 ATNKKYKYCWVWLGVKHSQVPMWOKKPLYLGAFLNPLAKNFTLNGKENLA
 WFLGGTGGRLKGDWATVREYVEALSVEIDVSGIRGNLLKFWFAQIAJNVDP
 KEANFNKYKGSALYMWGITDLSFRAYGAYSKPANDKLSDFTEFKDLGISAGF
 KSNFNKILKORLILISEFFFIYNKTKRLLFEFLVNPFGPGPIDETERTPPAD
 LSAQLEAANKSAEQRJAGAEKPKESKTSVERMSILRSVNALMSLADKLGA
 SSSSSSTSSADVSTTATPTPPPTFDYKTAQTDYDTFTSTSLADIQAALVS
 LQDAVTKIATQADETAETAEWETKNADAVKVAQITELAKVSDNQALDLSLGL
 TSPDLLQAKLLVANNKAEILLKEMDNFVPGKTPAIAQSLVDQDTATQIEKD
 GNARDFAYAGORAGAVENAKSNSISNIDSAAKATATQIAEAKQKFPSPILQ
 EASQWVZQAEKDLNKPQSDVFNFTVGGSKQSGISIRVMSLDDDAENETA
 SILMSGFQMHMFNTENPDSAAQELAAQRAAKAAGDDSAALADAQKALEAL
 GKAGQQQGIILALGQIASAAVWSVGLPQQVL"
 ORIGIN
 Query Match 99.2%; Score 1444; DB 1; Length 2148;
 Best Local Similarity 99.9%; Pred. No. 8.6e-299;
 Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 1 ATAAATCTTTAAAAACAGGCTCGCATTAATTATTAGTGAGAGCTTTTATTTTATTTT 60
 668 ATAAATCTTTAAAAACAGGCTCGCATTAATTATTAGTGAGAGCTTTTATTTTATTTT 727
 61 ATATAAACTAAAGATTTTATTATTTTGTAGTTTAA-TGGTTAACTTATTCGTC 119
 728 ATATAAACTAAAGATTTTATTATTTTGTAGTTTAA-TGGTTAACTTATTCGTC 787
 120 CAGTCTCTATAGAGCAACAGACGACACTCCCGCAGATCTTCTTCGTCACAGATTGG 179
 788 CAGTCTCTATAGAGCAACAGACGACACTCCCGCAGATCTTCTTCGTCACAGATTGG 847
 180 AGCGAGTGCAGCAAAATAAGAGTGCAGGAGCTCAAAGAAATAGCAGGTGCGGAAGCTAAGC 239
 848 AGCGAGTGCAGCAAAATAAGAGTGCAGGAGCTCAAAGAAATAGCAGGTGCGGAAGCTAAGC 907
 240 CTAAGAACTTAAGACCGATTCTGTAGAGGATGGAGCACTTCGCTTCTGAGTGAATG 299
 908 CTAAGAACTTAAGACCGATTCTGTAGAGGATGGAGCACTTCGCTTCTGAGTGAATG 967
 300 CTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGCTTCTACTA 359
 968 CTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGCTTCTACTA 1027
 360 GCAGATCTGCAGAGCTGGAATCAACGACAGCGACCGACCTTACGCTTCCTCCACCCACGT 419

1028 GCAGATCTGCAGAGGTGGACTCAACGACACGACCGCACCTACGCTCTCTCCACCCAGGT 1087
 420 TTGATGATTATAAGACTCAAGCGCAACACAGCTTTACGATACTATCTTTTACCTCAACATCAC 479
 1088 TTGATGATTATAAGACTCAAGCGCAACACAGCTTTACGATACTATCTTTTACCTCAACATCAC 1147
 480 TAGCTGACATACAGGCTGCTTTGGTGAAGCTTCAGAGTGTCTGTCACATAATAAAGGATA 539
 1148 TAGCTGACATACAGGCTGCTTTGGTGAAGCTTCAGAGTGTCTGTCACATAATAAAGGATA 1207
 540 CAGCGGCTACTGATAGGAAACGCAATCTCGGAGTGGGAAATAAGAAATGCGCGATG 599
 1208 CAGCGGCTACTGATAGGAAACGCAATCTCGGAGTGGGAAACTAAGAAATGCGCGATG 1267
 600 CAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGA 659
 1268 CAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGA 1327
 660 TTCTTGACTCTTTAGGTAAACGCTGACTCTCTTCGACCTTTACAGGCTGCTTCTTCTCAAT 719
 1328 TTCTTGACTCTTTAGGTAAACGCTGACTCTCTTCGACCTTTACAGGCTGCTTCTTCTCAAT 1387
 720 CTGTAGCAAAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTAG 779
 1388 CTGTAGCAAAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTAG 1447
 780 TCCAGGAAACCGCTGCAATTGCTCAATCTTTTAGTTGATCAGACAGATGCTTACAGCGA 839
 1448 TCCAGGAAACCGCTGCAATTGCTCAATCTTTTAGTTGATCAGACAGATGCTTACAGCGA 1507
 840 CACAGTAGAAGAAGATGGAAATGCGATAGGAGTGCATATTTTGCAGGACAGAACGCTA 899
 1508 CACAGTAGAAGAAGATGGAAATGCGATAGGAGTGCATATTTTGCAGGACAGAACGCTA 1567
 900 GTGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAG 959
 1568 GTGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAG 1627
 960 CAGCAATCGGTACTGTCTAAGACCAATAGCTGAAGCTCAGAAAAGTTCCCGACCTCTC 1019
 1628 CAGCAATCGGTACTGTCTAAGACCAATAGCTGAAGCTCAGAAAAGTTCCCGACCTCTC 1687
 1020 CAATTTCTCAGAAGCGGACAAATGGTATATACAGCTGGAAGATCTTTAAATAATCA 1079
 1688 CAATTTCTCAGAAGCGGACAAATGGTATATACAGCTGGAAGATCTTTAAATAATCA 1747
 1080 AACCTGCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAAC 1139
 1748 AACCTGCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAAC 1807
 1140 AAGAAAGTAGTATGTTAGTATTCGTGTTCCATGCTGTTAGATGATGCTGGAATAAGAGA 1199
 1808 AAGAAAGTAGTATGTTAGTATTCGTGTTCCATGCTGTTAGATGATGCTGGAATAAGAGA 1867
 1200 CCGCTTCATTTTGTGATGCTGGGTTCGTCAGATGATTCACATGTTCAATACCGAAAATC 1259
 1868 CCGCTTCATTTTGTGATGCTGGGTTCGTCAGATGATTCACATGTTCAATACCGAAAATC 1927
 1260 CTGATTCTCAAGCTGCCCAACAGAGCTCGCAGCAACAAGCTAGACGAGCGAAAGCGCTG 1319
 1928 CTGATTCTCAAGCTGCCCAACAGAGCTCGCAGCAACAAGCTAGACGAGCGAAAGCGCTG 1987
 1320 GAGATCAGAGTGTCTGTCAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGGGCTCTAG 1379
 1988 GAGATCAGAGTGTCTGTCAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGGGCTCTAG 2047
 1380 GTAAAGCTGGCAACAAAGGCGCATCTCAATCTTTTAGCAGATCGCTTCTGCTGCTG 1439
 2048 GTAAAGCTGGCAACAAAGGCGCATCTCAATCTTTTAGCAGATCGCTTCTGCTGCTG 2107
 1440 TTGTGAGCGCAGGAGT 1455
 2108 TTGTGAGCGCAGGAGT 2123


```
ORIGIN
Query Match 93.1%; Score 1355; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 1e-279;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGGTTAATCTTATGTCAGGTCCTATATAGACGAAACAGACGACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTTATGTCAGGTCCTATATAGACGAAACAGACGACACCTCCCGCAGAT 60
QY 161 CTTTCTGCTCAAGATTGGAGGCGAGTCGACGAAATAAGAGTCGGGAAGCTCAAGAAATA 220
Db 61 CTTTCTGCTCAAGATTGGAGGCGAGTCGACGAAATAAGAGTCGGGAAGCTCAAGAAATA 120
QY 221 GCAGGTGCGGAAGCTAAGCCCTAAAGAACTTAAGACCGATTCTGTAGAGCGATGGAGCATC 280
Db 121 GCAGGTGCGGAAGCTAAGCCCTAAGAACTTAAGACCGATTCTGTAGAGCGATGGAGCATC 180
QY 281 TTGGTTCTGCAAGTATGCTCATAGTCTGGCAGATGAAGCTGGGTATGCTTCTAGT 340
Db 181 TTGGTTCTGCAAGTATGCTCATAGTCTGGCAGATGAAGCTGGGTATGCTTCTAGT 240
QY 341 AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGACGCTCAACGACGACGACCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGACGCTCAACGACGACGACCGACCT 300
QY 401 ACGCTCTCTCCAGCCAGCTTTGATGATTATAAGACTCAAGCGCAAAACAGCTTACGATPACT 460
Db 301 ACGCTCTCTCCAGCCAGCTTTGATGATTATAAGACTCAAGCGCAAAACAGCTTACGATPACT 360
QY 461 ATCTTTACCTCAACATCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 520
Db 361 ATCTTTACCTCAACATCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 521 GTCACTAATATAAGGATACAGCGGCTCTGATGAGGAAACCGCAATCGCTGGGAGTGG 580
Db 421 GTCACTAATATAAGGATACAGCGGCTCTGATGAGGAAACCGCAATCGCTGGGAGTGG 480
QY 581 GAACTAAGAAATGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
Db 481 GAACTAAGAAATGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 641 GCTTCGATTAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 700
Db 541 GCTTCGATTAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 701 CAGGCTGCTCTTCTCCAACTGTAGCAAAACAAATAACAGCAGCTGAGCTTCTTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAACTGTAGCAAAACAAATAACAGCAGCTGAGCTTCTTAAAGAG 660
QY 761 ATGCAAGATAACCCAGTAGTCCCAGGGAACACGCTGCAATTTGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTAGTCCCAGGGAACACGCTGCAATTTGCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATGCTACAGGACACAGATAGAGAAATGGAATGCAATGAGGATGCAAT 880
Db 721 CAGACAGATGCTACAGGACACAGATAGAGAAATGGAATGCAATGAGGATGCAAT 780
QY 881 TTTTCAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 940
Db 781 TTTTCAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAATACTAATAACAGTATAAGC 840
QY 941 AACATAGATTACGTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 1000
Db 841 AACATAGATTACGTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 900
QY 1001 AAAAAGTCCCGGACTCTCCAAATTTCTCAGAACGGAACAAATGGTAAATACAGGCTGAG 1060
Db 901 AAAAAGTCCCGGACTCTCCAAATTTCTCAGAACGGAACAAATGGTAAATACAGGCTGAG 960
QY 1061 AAAGATCTTAAATAATCAAACTGCGAGATGGTCTGATGTTCCAAATCCAGGAACCTACA 1120
```

```
Db 961 AAAGATCTTAAATAATCAAACTGCGAGATGGTCTGATGTTCCAAATCCAGGAACCTACA 1020
QY 1121 GTTGGAGGCTCCAAAGCAACAAAGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGGCTCCAAAGCAACAAAGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTTA 1080
QY 1181 GATGATGCTGAAATAGAGACCGCTTCCATTTGATGCTCTGGGTTTCTGTCAGATGATTCAC 1240
Db 1081 GATGATGCTGAAATAGAGACCGCTTCCATTTGATGCTCTGGGTTTCTGTCAGATGATTCAC 1140
QY 1241 ATGTTCAATACGAAATCTCTGATTTCAAGCTGCCCCAACAGGAGCTCGCAGCACAGCT 1300
Db 1141 ATGTTCAATACGAAATCTCTGATTTCAAGCTGCCCCAACAGGAGCTCGCAGCACAGCT 1200
QY 1301 AGAGCAGCAAAAGCCCTGGAGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAA 1360
Db 1201 AGAGCAGCAAAAGCCCTGGAGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAA 1260
QY 1361 GCTTTAGACCGGCTCTAGTAAAGCTGGGCAACAAAGGCGATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGACCGGCTCTAGTAAAGCTGGGCAACAAAGGCGATCTCAATGCTTTAGGA 1320
QY 1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1455
Db 1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1355

RESULT 11
AX349491
LOCUS AX349491 1956 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 14 from Patent WO0202606.
ACCESSION AX349491
VERSION AX349491.1 GI:18615352
KEYWORDS
SOURCE
ORGANISM Chlamydomophila pneumoniae
Chlamydomophila pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE
1 Ratti, G. and Grandi, G.
AUTHORS Immunisation against Chlamydia pneumoniae
TITLE Patent: WO 0202606-A 14 10-JAN-2002;
JOURNAL Chiron S.p.A. (IT)
FEATURES
source
location/qualifiers
1..1956
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
ORIGIN
Query Match 93.1%; Score 1355; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 1e-279;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGGTTAATCTTATGTCAGGTCCTATATAGACGAAACAGACGACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTTATGTCAGGTCCTATATAGACGAAACAGACGACACCTCCCGCAGAT 60
QY 161 CTTTCTGCTCAAGATTGGAGGCGAGTCGACGAAATAAGAGTCGGGAAGCTCAAGAAATA 220
Db 61 CTTTCTGCTCAAGATTGGAGGCGAGTCGACGAAATAAGAGTCGGGAAGCTCAAGAAATA 120
QY 221 GCAGGTGCGGAAGCTAAGCCCTAAAGAACTTAAGACCGATTCTGTAGAGCGATGGAGCATC 280
Db 121 GCAGGTGCGGAAGCTAAGCCCTAAGAACTTAAGACCGATTCTGTAGAGCGATGGAGCATC 180
QY 281 TTGGTTCTGCAAGTATGCTCATAGTCTGGCAGATGAAGCTGGGTATGCTTCTAGT 340
Db 181 TTGGTTCTGCAAGTATGCTCATAGTCTGGCAGATGAAGCTGGGTATGCTTCTAGT 240
QY 341 AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGACGCTCAACGACGACGACCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGACGCTCAACGACGACGACCGACCT 300
QY 401 ACGCTCTCTCCAGCCAGCTTTGATGATTATAAGACTCAAGCGCAAAACAGCTTACGATPACT 460
Db 301 ACGCTCTCTCCAGCCAGCTTTGATGATTATAAGACTCAAGCGCAAAACAGCTTACGATPACT 360
QY 461 ATCTTTACCTCAACATCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 520
Db 361 ATCTTTACCTCAACATCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 521 GTCACTAATATAAGGATACAGCGGCTCTGATGAGGAAACCGCAATCGCTGGGAGTGG 580
Db 421 GTCACTAATATAAGGATACAGCGGCTCTGATGAGGAAACCGCAATCGCTGGGAGTGG 480
QY 581 GAACTAAGAAATGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
Db 481 GAACTAAGAAATGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 641 GCTTCGATTAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 700
Db 541 GCTTCGATTAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 701 CAGGCTGCTCTTCTCCAACTGTAGCAAAACAAATAACAGCAGCTGAGCTTCTTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAACTGTAGCAAAACAAATAACAGCAGCTGAGCTTCTTAAAGAG 660
QY 761 ATGCAAGATAACCCAGTAGTCCCAGGGAACACGCTGCAATTTGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTAGTCCCAGGGAACACGCTGCAATTTGCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATGCTACAGGACACAGATAGAGAAATGGAATGCAATGAGGATGCAAT 880
Db 721 CAGACAGATGCTACAGGACACAGATAGAGAAATGGAATGCAATGAGGATGCAAT 780
QY 881 TTTTCAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 940
Db 781 TTTTCAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAATACTAATAACAGTATAAGC 840
QY 941 AACATAGATTACGTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 1000
Db 841 AACATAGATTACGTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 900
QY 1001 AAAAAGTCCCGGACTCTCCAAATTTCTCAGAACGGAACAAATGGTAAATACAGGCTGAG 1060
Db 901 AAAAAGTCCCGGACTCTCCAAATTTCTCAGAACGGAACAAATGGTAAATACAGGCTGAG 960
QY 1061 AAAGATCTTAAATAATCAAACTGCGAGATGGTCTGATGTTCCAAATCCAGGAACCTACA 1120
```


401 ACGCTCTCCACCCAGTTTGTATATATAGACTCAGCGCAACAGCTTACGATCT 460
Db |||
301 ACGCTCTCCACCCAGTTTGTATATATAGACTCAGCGCAACAGCTTACGATCT 360
Qy |||
461 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGT 520
Db |||
361 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGT 420
Qy |||
521 GTCACTAATATAAGGATACAGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTG 580
Db |||
421 GTCACTAATATAAGGATACAGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTG 480
Qy |||
581 GAACTTAAGATGCGGATGAGTTAAGTTGGCGCAAAATACAGAAATAGCGAAATAT 640
Db |||
481 GAACTTAAGATGCGGATGAGTTAAGTTGGCGCAAAATACAGAAATAGCGAAATAT 540
Qy |||
641 GCTTCGATTAACCAAGCGATCTTGACTCTTTAGGTAAGCTGACTCTTCCTCGACCTCTTA 700
Db |||
541 GCTTCGATTAACCAAGCGATCTTGACTCTTTAGGTAAGCTGACTCTTCCTCGACCTCTTA 600
Qy |||
701 CAGGCTGCTCTTCTCAATCTGTAGCAACAAATACAAAGCGCTGAGCTTCTTAAGAG 760
Db |||
601 CAGGCTGCTCTTCTCAATCTGTAGCAACAAATACAAAGCGCTGAGCTTCTTAAGAG 660
Qy |||
761 ATGCAAGATAACCCAGTAGTCCCGGAGAAACCGCTGCAATGCTCAATCTTTAGTTGAT 820
Db |||
661 ATGCAAGATAACCCAGTAGTCCCGGAGAAACCGCTGCAATGCTCAATCTTTAGTTGAT 720
Qy |||
821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATAGGAGATGATAT 880
Db |||
721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATAGGAGATGATAT 780
Qy |||
881 TTTGCGAGGACAGAACCTAGTGAGCTGTAGAAATGCTAAATCTTAATACAGTATAAG 940
Db |||
781 TTTGCGAGGACAGAACCTAGTGAGCTGTAGAAATGCTAAATCTTAATACAGTATAAG 840
Qy |||
941 AACATAGATTACGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 1000
Db |||
841 AACATAGATTACGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 900
Qy |||
1001 AAAAGTTCCCGACTCTCAATCTTCAAGAAAGCGGAAACAAATGTAATACAGGCTGAG 1060
Db |||
901 AAAAGTTCCCGACTCTCAATCTTCAAGAAAGCGGAAACAAATGTAATACAGGCTGAG 960
Qy |||
1061 AAGATCTTAAATATCAAACTGAGATGCTGCTGATGCTTCAATCCGAGAACTACA 1120
Db |||
961 AAGATCTTAAATATCAAACTGAGATGCTGCTGATGCTTCAATCCGAGAACTACA 1020
Qy |||
1121 GTTGGAGGCTCCAAGCAACAAAGGAAGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTA 1180
Db |||
1021 GTTGGAGGCTCCAAGCAACAAAGGAAGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTA 1080
Qy |||
1181 GATGATGCTGAAATGAGACCGCTTCGATTTGATGCTGGGTTTGTGATGATTTAC 1240
Db |||
1081 GATGATGCTGAAATGAGACCGCTTCGATTTGATGCTGGGTTTGTGATGATTTAC 1140
Qy |||
1241 ATGTTCAATACGGAATCTGATTTCTCAAGCTGCCCAACAGAGGCTCGCAGACCAAGCT 1300
Db |||
1141 ATGTTCAATACGGAATCTGATTTCTCAAGCTGCCCAACAGAGGCTCGCAGACCAAGCT 1200
Qy |||
1301 AGAGCAGGAAACCGCTGAGATGACAGTGTCTGCTGAGCGCTGCGAGATGCTCAGAAA 1360
Db |||
1201 AGAGCAGGAAACCGCTGAGATGACAGTGTCTGCTGAGCGCTGCGAGATGCTCAGAAA 1260
Qy |||
1361 GCTTTAGAGCGGCTCTAGGTAAGCTGGCAACACAGGCGATCTCAATGCTTTAGGA 1420
Db |||
1261 GCTTTAGAGCGGCTCTAGGTAAGCTGGCAACACAGGCGATCTCAATGCTTTAGGA 1320
Qy |||
1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1455
Db |||
1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1355

RESULT 12
AX361992 1956 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 385 from Patent WO0208267.
DEFINITION AX361992
ACCESSION AX361992
KEYWORDS AX361992.1 GI:18694447
SOURCE Chlamydomophila pneumoniae
ORGANISM Chlamydomophila pneumoniae
Bacteria; Chlamydia; Chlamydiales; Chlamydiaceae; Chlamydomophila.
1
REFERENCE
AUTHORS Fling, S.P., Skeiky, Y.A., Probst, P. and Bhatia, A.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0208267-A 385 31-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..1956
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
ORIGIN
Query Match 93.1%; Score 1355; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 1e-279;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 ATGGTTAATCTTATGGTTCAGGTCCTATAGACGAAACAGACGACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTTATGGTTCAGGTCCTATAGACGAAACAGACGACACCTCCCGCAGAT 60
QY 161 CTTTCCTGCTCAAGGATTCGAGCGAGTCAGCAATATAGAGTCGCGAAGCTCAAAAGATA 220
Db 61 CTTTCCTGCTCAAGGATTCGAGCGAGTCAGCAATATAGAGTCGCGAAGCTCAAAAGATA 120
QY 221 GCAGTGGCGAGCTTAAGCCTAAAGAACTTAAGACCGATTCGTAGAGCGATGAGAGATC 280
Db 121 GCAGTGGCGAGCTTAAGCCTAAAGAACTTAAGACCGATTCGTAGAGCGATGAGAGATC 180
QY 281 TTGCGTTCTGAGTGAATGCTCTCATGATCTGGCAGATAGCTGGTATTTCTTCTAGT 340
Db 181 TTGCGTTCTGAGTGAATGCTCTCATGATCTGGCAGATAGCTGGTATTTCTTCTAGT 240
QY 341 AACAGCTGCTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGCGACCCACCT 400
Db 241 AACAGCTGCTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGCGACCCACCT 300
QY 401 AGCCTCTCTCAACCGCTTTGATGATTAAGATCAAGCGCAACAGCTTACGATCT 460
Db 301 AGCCTCTCTCAACCGCTTTGATGATTAAGATCAAGCGCAACAGCTTACGATCT 360
QY 461 ATCTTTACTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGT 520
Db 361 ATCTTTACTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGT 420
QY 521 GTCACTAATATAAGGATACAGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTG 580
Db 421 GTCACTAATATAAGGATACAGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTG 480
QY 581 GAACTTAAGATGCGGATGAGTTAAGTTGGCGCAAAATACAGAAATAGCGAAATAT 640
Db 481 GAACTTAAGATGCGGATGAGTTAAGTTGGCGCAAAATACAGAAATAGCGAAATAT 540
QY 641 GCTTCGATTAACCAAGCGATCTTGACTCTTTAGGTAAGCTGACTCTTCCTCGACCTCTTA 700
Db 541 GCTTCGATTAACCAAGCGATCTTGACTCTTTAGGTAAGCTGACTCTTCCTCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCTCAATCTGTAGCAACAAATACAAAGCGCTGAGCTTCTTAAGAG 760
Db 601 CAGGCTGCTCTTCTCAATCTGTAGCAACAAATACAAAGCGCTGAGCTTCTTAAGAG 660
QY 761 ATGCAAGATAACCCAGTAGTCCCGGAGAAACCGCTGCAATGCTCAATCTTTAGTTGAT 820


```
Db 661 ATGCAAGATACCCAGTAGTCCAGGGAACGCTGCAATTGCTCAATCTTATGTTGAT 720
Qy 821 CAGACAGATGCTACAGGACACAGATAGAGAAAGATGGAATGCGATGAGGATGATAT 880
Db 721 CAGACAGATGCTACAGGACACAGATAGAGAAAGATGGAATGCGATGAGGATGATAT 780
Qy 881 TTTCAGGACAGAACGCTAGTGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAGC 940
Db 781 TTTCAGGACAGAACGCTAGTGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAGC 840
Qy 941 AACATAGATTCAGCTAAAGCAGCAATCGCTACTCTTAAGACACAAAATAGCTGAGCTCAG 1000
Db 841 AACATAGATTCAGCTAAAGCAGCAATCGCTACTCTTAAGACACAAAATAGCTGAGCTCAG 900
Qy 1001 AAAAAGTCCCGGACTCTCCAAATCTTCAAGACGGGAACAAATGGTAAATACAGGCTGAG 1060
Db 901 AAAAAGTCCCGGACTCTCCAAATCTTCAAGACGGGAACAAATGGTAAATACAGGCTGAG 960
Qy 1061 AAAGATCTTAAATAATCAAACTGCGAGATGGTTCTGATGTTCCAAATCCAGGAACCTACA 1120
Db 961 AAAGATCTTAAATAATCAAACTGCGAGATGGTTCTGATGTTCCAAATCCAGGAACCTACA 1020
Qy 1121 GTTGGAGGCTCCAAAGCAACAGGAAGTAGTATGTTGTTTCCATGCTCTTA 1180
Db 1021 GTTGGAGGCTCCAAAGCAACAGGAAGTAGTATGTTGTTTCCATGCTCTTA 1080
Qy 1181 GATGATCTGAAATGAGCGCTTCCATTTGATGTTCTGGGTTCTGATGATGATTCAC 1240
Db 1081 GATGATCTGAAATGAGCGCTTCCATTTGATGTTCTGGGTTCTGATGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATCTCTGATTCCTCAAGCTGCCCAACAGGAGCTCGGACACAGCT 1300
Db 1141 ATGTTCAATACGGAATCTCTGATTCCTCAAGCTGCCCAACAGGAGCTCGGACACAGCT 1200
Qy 1301 AGACGCGGAAGCGCTGAGATGACAGTCTGCTGACGCTGCGGATGCTCAGAAA 1360
Db 1201 AGACGCGGAAGCGCTGAGATGACAGTCTGCTGACGCTGCGGATGCTCAGAAA 1260
Qy 1361 GCTTTAGAGCGCTCTAGTAAAGCTGGGCAACAGCGGATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGCTCTAGTAAAGCTGGGCAACAGCGGATCTCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTCTGCTGCTGTTGTCAGCGCAGGAGT 1455
Db 1321 CAGATCGCTCTGCTGCTGTTGTCAGCGCAGGAGT 1355
```

```
RESULT 13
AX045133 1852 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 3 from Patent WO0066739.
DEFINITION AX045133
ACCESSION AX045133
VERSION AX045133.1 GI:11343732
KEYWORDS Chlamydomophila pneumoniae
SOURCE Chlamydomophila pneumoniae
ORGANISM Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.
REFERENCE 1
AUTHORS Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
TITLE 3(chlamydia) antigens and corresponding dna fragments and uses thereof
JOURNAL Patent: WO 0066739-A 3 09-NOV-2000;
Aventis Pasteur Limited (CA)
FEATURES
source Location/Qualifiers
1. 1852
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
1. 1752
/note="unnamed protein product"
/codon_start=1
/transl_table=11
CDS
```

```
/protein_id="CAC17378.1"
/db_xref="GI:11343733"
/translation="MSLADKLGIASSSSSSSTSRSDVSTTATATPTPPPTDDYKT
QAQATYDTFTSTLADIQAALVQLQDAVTNIDKTADEETAEIAAEWETKADAVKV
GAQITELAKYASDNQAILDSGLKLTGDLLOALLOSVANNKAAELKEMQONPVVP
KPTAATQSLVDOTDATATIEKGNARIDAVFAGNAGAVENAKSNNSISINIDSAK
AATATAQTQIAEAKKPPSPILQEAQWVIOAEKDLKNIKPADGSDVNPPTVGGG
KQGSISIGSRVMSLDDAENETASILMSGFROIMHMENTENPDSQAQBELAAQARA
AKAAGDSSAAALADAKALEAALGKAGQOQILNALGOTASAAVSVAGVPPAAASI
GSSVGLYKTSKSTGSDYKTSAGYDAYKSDINAYGRANDATRDVINNVSPALTR
SVPRATEARGPEKTPQALARVSGNSRTGLDVYSOVSALOSVMQITQSPQANBEI
ROKLTSAVTKPPOFGYPYVQLSNDSTQKFLAKLESFABGSRTAABEIKALSFTNSLF
IQQVNIYNSLSISGLYQ"
ORIGIN
Query Match 79.1%; Score 1151; DB 6; Length 1852;
Best Local Similarity 100.0%; Pred. No. 5.1e-236;
Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 305 ATGAGTCTGCGACATAAGCTGGGTATGCTTCTAGTAAACAGCTGCTTCTTACTAGCAGA 364
Db 1 ATGAGTCTGCGACATAAGCTGGGTATGCTTCTAGTAAACAGCTGCTTCTTACTAGCAGA 60
Qy 365 TCTGCAGACGCTGACTCAACGACGACGACGACCTACGCTCCTCCACCCACGTTTGTAT 424
Db 61 TCTGCAGACGCTGACTCAACGACGACGACGACCTACGCTCCTCCACCCACGTTTGTAT 120
Qy 425 GATTATAAGACTCAACGCGCAACAGCTTACGATPACTATCTTTTACCTCAACATCACTAGCT 484
Db 121 GATTATAAGACTCAACGCGCAACAGCTTACGATPACTATCTTTTACCTCAACATCACTAGCT 180
Qy 485 GACATACAGCTGCTTGTGTGAGCTCCAGGATGCTCTCACTAATATAAGGATACAGCG 544
Db 181 GACATACAGCTGCTTGTGTGAGCTCCAGGATGCTCTCACTAATATAAGGATACAGCG 240
Qy 545 GCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCGGATGAGATT 604
Db 241 GCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCGGATGAGATT 300
Qy 605 AAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATTCCT 664
Db 301 AAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATTCCT 360
Qy 665 GACTCTTTAGGTAAAGCTGACTTCCCTTCGAGCTCTTACAGGCTGCTCTTCCCAATCTGTA 724
Db 361 GACTCTTTAGGTAAAGCTGACTTCCCTTCGAGCTCTTACAGGCTGCTCTTCCCAATCTGTA 420
Qy 725 GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 784
Db 421 GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Qy 785 GGGAAAACCGCTGCAATTCGTCATCTTTAGTTGATCAGACAGATGCTACAGGACACAG 844
Db 481 GGGAAAACCGCTGCAATTCGTCATCTTTAGTTGATCAGACAGATGCTACAGGACACAG 540
Qy 845 ATAGAGAAAGATGGAATCGGATGAGGATGATATTTTTCAGGACAGACGCTAGTGG 904
Db 541 ATAGAGAAAGATGGAATCGGATGAGGATGATATTTTTCAGGACAGACGCTAGTGG 600
Qy 905 GCTGTAGAAAATGCTAAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGCAGA 964
Db 601 GCTGTAGAAAATGCTAAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGCAGA 660
Qy 965 ATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAATT 1024
Db 661 ATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAATT 720
Qy 1025 CTTCAAGAGCGGAACAAATGGTATACAGCTGAGGAGATCTTAAATATCAACCT 1084
Db 721 CTTCAAGAGCGGAACAAATGGTATACAGCTGAGGAGATCTTAAATATCAACCT 780
Qy 1085 GCAGATGTTCTGATGTTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAAGGA 1144
```

```
Db      781  GCAGATGGTTCGATGTTCCAAATCCAGGACCTACAGTTGGAGGCTCCAGGACCAAGGA  840
QY      1145  AGTAGTATGGTAGTATTCGTGTTTCCATGCTGTAGATGATGCTGAAATGAGACCGCT  1204
Db      841  AGTAGTATGGTAGTATTCGTGTTTCCATGCTGTAGATGATGCTGAAATGAGACCGCT  900
QY      1205  TCCATTTGATGCTCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATCTCTGAT  1264
Db      901  TCCATTTGATGCTCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATCTCTGAT  960
QY      1265  TCTCAAGCTGCCCAACAGGAGCTCGCAGACAAGCTAGACGACGCAAAAGCCGCTGAGAT  1324
Db      961  TCTCAAGCTGCCCAACAGGAGCTCGCAGACAAGCTAGACGACGCAAAAGCCGCTGAGAT  1020
QY      1325  GACAGTCTGCTGCTGAGCGCTGCGAGATGCTCAGAAAAGCTTTAGAACGCGCTCTAGTAAA  1384
Db      1021  GACAGTCTGCTGCTGAGCGCTGCGAGATGCTCAGAAAAGCTTTAGAACGCGCTCTAGTAAA  1080
QY      1385  GCTGGGCAACACAGGAGCTCACTCAATGCTTTAGGACAGATCGTCTGCTGCTGTGTG  1444
Db      1081  GCTGGGCAACACAGGAGCTCACTCAATGCTTTAGGACAGATCGTCTGCTGCTGTGTG  1140
QY      1445  AGCGCAGGAGT  1455
Db      1141  AGCGCAGGAGT  1151
```

```
RESULT 14
AEO16994
LOCUS      304769 bp      DNA      linear      BCT 15-AUG-2003
DEFINITION Chlamydomophila caviae GPIC section 1 of 4 of the complete genome.
ACCESSION AEO16994 AEO15925
VERSION AEO16994.1 GI:29834116
KEYWORDS
SOURCE
ORGANISM Chlamydomophila caviae GPIC
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE 1 (bases 1 to 304769)
AUTHORS Read,T.D., Myers,G.S., Brunham,R.C., Nelson,W.C., Paulsen,I.T.,
Heidelberg,J., Holtzapple,E., Khouri,H., Fedorova,N.B., Carty,H.A.,
Umayam,L.A., Haft,D.H., Peterson,J., Beanan,M.J., White,O.,
Salzberg,S.L., Hsia,R.C., McClarty,G., Rank,R.G., Bavoil,P. and
Fraser,C.M.
TITLE Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
examining the role of niche-specific genes in the evolution of the
Chlamydiaceae
JOURNAL Nucleic Acids Res. 31 (8), 2134-2147 (2003)
MEDLINE 2269155
PUBMED 12682364
REFERENCE 2 (bases 1 to 304769)
AUTHORS Read,T., Myers,G., Brunham,R., Nelson,W., Paulsen,I.,
Heidelberg,J., Holtzapple,E., Khouri,H., Fedorova,N., Carty,H.,
Umayam,L., Haft,D., Peterson,J., Beanan,M., White,O., Salzberg,S.,
Hsia,R.-C., McClarty,G., Rank,R., Bavoil,P. and Fraser,C.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
location/Qualifiers
1. 304769
/organism="Chlamydomophila caviae GPIC"
/mol_type="genomic DNA"
/isolate="GPIC"
/db_xref="taxon:227941"
57..1052
/gene="hemB"
/locus_tag="CCA00001"
57..1052
/gene="hemB"
/locus_tag="CCA00001"
/notes="similar to GB:L16784, SP:P33395, and PID:290384;
identified by sequence similarity; putative"
/codon_start=1
```

```
/transl_table=11
/product="delta-aminolevulinic acid dehydratase"
/protein_id="AAP04754.1"
/db_xref="GI:29834117"
/translation="MSSIALLRPRNRKTAIRDLVAETALLPDFQIFCPFFVKEGKN
ICEETESLPVGRWSIDLLEKEIRLCGLKAVILFPVIPSDLKDAYGSSYNPKNI
LCKSYEVKAPPDLCVTSIDIALDPYTHGDMQCEVLNDESVRIFGNATIAHAE
MGADIVASDMMGDRVAHRAKLDOSGWTNLIISYSVKYASLLNRPDRDALGSLQS
GKNRYQNPKNVLEALLECEDEGADMLMKPAGLYLDVHLHVRNLSITLPLAAYQ
VSGEYALTAADAKYGLNKEAMFYESLIAIKRAGADMTISYATPLILELVSSKT"
complement (1068..2480)
/gene="nqrA"
/locus_tag="CCA00002"
complement (1068..2480)
/gene="nqrA"
/locus_tag="CCA00002"
/notes="similar to SP:Q56586, GB:L16784, SP:P33395, and
PID:290384; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, subunit A"
/protein_id="AAP04755.1"
/db_xref="GI:29834118"
/translation="MKIAITRGLDLQLQSPKESGFLKRIDPALVSVDLRPSALALK
LKVEQDVISGSPTAENPFGVITSPVGIKVEIRRKESLDDVVIKTPGONL
TEYSYDLKLSRAELLEIFKKEGLFALFKQRPDIPALPTOTPRDFINLADNRPTP
STEKQLAVFSREESGFYVNGRAVAKLFGLCPHIVSTDRLALEPEKDKSTAHHLKI
TGYPGSPSTHIYIATINKEIKVITLISFOEVLTIHGLFKGLINQEVVALAGSG
LKSLRRYVITTKGANFESLLPLQDISDVSLSIGDPLTGRCLRESPLCLGMRDSTI
SVLPNPKROAFNLRGLNKPTHTRTYLSGLFKKHTYMDPTNLHGETRPIIDTEI
YDKVMLKIPVVPVLIKAVITKDFELACMLGFLVCPEDFALPTFDPSKTEMLTIKD
ALKDYAKETGILNLEVEDKE"
complement (2502..2936)
/locus_tag="CCA00003"
complement (2502..2936)
/locus_tag="CCA00003"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP04756.1"
/db_xref="GI:29834119"
/translation="MNSNAOKIIDSIKOILTYNIDFPSPGSSSDSEVDYELM
EKTQKIOELDKRSODIILQOTGMTQEOMFANPDNPFSEWLALENIRSSCDAYK
ETEAALNVTKDLGSETITSSNNKTKTSSKKNKKKNKNIPL"
3048..5201
/gene="greA"
/locus_tag="CCA00004"
3048..5201
/gene="greA"
/locus_tag="CCA00004"
/notes="similar to GB:L10328, SP:P04846, GB:M12163,
PID:L46649, PID:290510, GB:U00096, and PID:L790093;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="transcription elongation factor GreA"
/protein_id="AAP04757.1"
/db_xref="GI:29834120"
/translation="MDYLEKLQVLIDEBQPSFFNLWEEYCFNDVVVGAELVQILEKV
KHSLAPLFGKIADTVLPWMI PDGKEDRVLQI LDVQNTNAPFYDAADIDYVNRK
YAGRENFENALRVGLRGREFQYSLRFDLMLHSEGNFVPHSGVGWGVMSYFL
QOKVLI EFEGVVMKADISFETAKSLIPLEKDHFLSRFGPDGDFGSAKHPSPVIE
MLLKDLGPKTAKEIKDELVDLVIPEAEWNWQATKTKIKNI CITTPTKIDYRYN
PGDSLITQLETRLSQIEDTSKLVIEYQIFRDLHSELKPKENREIVIKALQTLUSVEG
NQSLEIQRDLLISDFLGEKSVLDSFSEELIVANNISIVVLQKRAFLMLIK
HSPVMENVMKIFLSTTSPSRELTFFKVRGEDSVYKVKIEKLLCEAHPMWYEVFT
WFELKGSDDGIFKPNDKIERLLEAALVPMYHVASTPOKELGKITYTLVSORYL
AVRNMEGAPLSYLKEILLSTKCSQFSSDLVLQSLAEVVDPLKNTVTVVVEDIL
WTTPBSFTKMNKQLQSLSGKEMVNDAKEIEDARVLGDLRENSFKPALEKARQSEI
RVLSEVNRARILKDIVFTDKVGVGVCKVLEDENGNLITYSILGPDADPDNYLILS
KSKLAQEMLDKSVGAAVQFGKKYKISRISQIWDE"
5353..6543
```

gene

```

/locus_tag="CCA00005"
5353.._8543
/locus_tag="CCA00005"
/notes="similar to SP:P04693, GB:L10328, SP:P04846,
GB:M12163, PID:146649, PID:290510, GB:U00096, and
PID:1790093; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="amino acid biosynthesis aminotransferase, class
1"
/protein_id="AAP04758.1"
/db_xref="GI:29834121"
/transl_table=11
/translation="MSFENLPTRFSDSLGQKLFEDEREKVNLIYGVEDPSKA
YGGFSVKAQPLFEDENKGLYPLISGLSSNOEMKLVFQDVNNSFVQAALGCT
GALHLGAKLFVAYTSKYLIPETQWGNHVRIPQAGLEVILKPIYSSKSLFDEM
LSVLAAPNSLVLIQCCNPTGMDLNMWRLAEVKNKHLQPLFFEDTAYLGFGRG
IKEDRRPELIEFVGNPVFVADASKNFSLYGERVGFVAVHSKITEDLDKSLCEEK
ARGEYSPFRGAKFVITLSGLKEMWLSLSDTIRSLGKTRARLVQAMNHHVGS
FDPLISQKGFPGFPGFVEQLFLRSEKGIYVTSGARFNLGITDKNIDHVSQSPAEA
YQLS"
6839.._7861
/locus_tag="CCA00006"
6839.._7861
/locus_tag="CCA00006"
/notes="similar to GB:L10328, SP:P04846, GB:M12163,
PID:146649, PID:290510, GB:U00096, and PID:1790093;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP04759.1"
/db_xref="GI:29834122"
/translation="MAGSVVRHKKRLVYLVILSVLWLSLPRFFENIKQHFV
SLYRVFKRSVPISDIQVENVILKRIALIEERLQALFEIYANIPVPFSEILSP
YFRKIDSRVIRPAHWSSQWNVNGKRIQKNSPVLQKLVGLVYDYGQOQSRI
RLITDVGQSPVIAVRGIGIOMLVKDOIQGLTKHLERLSDVILEKRYKVDQLEEL
SSIIQCSDDNTLLGGLSGGGPLMKDETLVHGEGFCFSDGKGLRVGDILITGLD
GVFPPLFVAEITKVCQREGACSYKIEAKSLAEPLTNLSVLILPAVEFNDRPDI
FGLLWD"
complement(7821.._10958)
/gene="recB"
/locus_tag="CCA00007"
complement(7821.._10958)
/gene="recB"
/locus_tag="CCA00007"
/notes="identified by match to TIGR protein family HMM
TIGR01866"
/codon_start=1
/transl_table=11
/product="exodeoxyribonuclease V, beta subunit"
/protein_id="AAP04760.1"
/db_xref="GI:29834123"
/translation="MKPDIENPQTSIRGKYFLASAGTGKFTIEQIVLRALLEGSV
SHVENLAVTFNATNEKURIQNLKQALCQIKSVLEDPKSLPPYLKDTSNVLL
YQVNAVALTDRAVATIHGFCNYVLQHPQMOTQKNAALTHSQAVFHHIKKYL
QDLWSVLFPPQVFLAAKXNSKHTAFLTDKLSVYLTQTFHLPTSTTDLARS
WHNSIRSKIQDIPKEDRQILLOHKDGFKKQFPIDEDRLFVECLVASETSVOLFSF
SKIAETFPKRLPTEYSKAFSHIEETLSLYTEQFCNBIENLILQVQLYLNKH
YTMLSPELIALBEILLSRSEIEIETLRKFOLVILIDFQDTDKQWILFKLFA
NDAFSGSLFDGPQSYIYERNADLATYLRKASFPKSSQLHILNNYRSTQAUMEAV
NILFCQFPFQIDPSYIEYHPLCFQSEHFENSHAPIHFFSYDDISDQAAWISHT
ASYLQATYHILPLGRMIALVSDSAQAFDLITHCSIPVAFSKNSIFHLPTVLLTLAW
EAILHVENYKIORVLLSLFRNLSDTIEKKELSYTFPSLSRVSDYGLANFYHF
MTLQCEALLKTPQDLSQEMERLCAYLDTVSSHQQLLYLQVSEIGRNEESLSFS
SYSETEILKTIITHASKLEYDVVFCGLDKSKKNSASEWIREMYVACTRANKQLF
IPQNSNRCSNALTNYKVHVSYESILDJAQLTKGHPRLFSITPSETSDPV
YKIIIPSTFISIPAKQIFSGSVKLAFDNEVLDNPVAFVAFVLPGRKGTGII
KHILEINSPFKISPSIVTFVKNTHLEGEYELISQKLVTFSSPLSPAFESRS
LQDICPSKIACESPFLSNQOLWQVLDLPEYNDKXYIIDWKTSGEISLSEKE
NLLNVIKQNLDYQAIYTHAAKRFLOQFDIVSDVMGFIFIRGMDPDEFGFLCPNQ
KISNPTITRKCPAYH"
complement(10945.._14010)
/gene="recC"

```

```

/locus_tag="CCA00008"
complement(10945.._14010)
/gene="recC"
/locus_tag="CCA00008"
/notes="similar to GB:L125079, SP:P42822, and PID:529423;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="exodeoxyribonuclease V, gamma subunit"
/protein_id="AAP04761.1"
/db_xref="GI:29834124"
Query Match      14.0%; Score 204.2; DB 1; Length 304769;
Best Local Similarity 50.4%; Fred. No. 1.9e-33;
Matches 746; Conservative 0; Mismatches 633; Indels 102; Gaps 6;
QY 1 ATAAAATCTTTAAAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTT 60
DB 21625 ATAAAATCTTTAAAAAGAGCTCACATCTAATTAATAAAGATGTGAGCTTTT 21684
QY 61 ATAATAAACTTAAAGATTTTATTTATTTTGTAGTTTTTATGGTTAAICCTATTGGTCC 120
DB 21685 TTATAATTTAAGAAAAGATTTTATTAGTTTTTTTATTATTATGTTAATCTCTGT 21738
QY 121 AGGTCCTATAGACGAAACAGAACCCACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 180
DB 21739 CGCCCTATAGATGAATCAAAAACATTGCTCTCGCAGACTTACTTCTTAGGTATGCA 21798
QY 181 GCGAGTGCAGCAATAAGAGTGGGAGGCTCAAGAATAGCAGGTGGGAAAGCTTAAGCC 240
DB 21799 GCGAGCGCAGCAATACTGTAGCTCAGAAGCTCAATCGATAACCGAATTCGAGGCAATC 21858
QY 241 TAAAGATCTAAGAC-----CGATTCTGTAGAGGATGAGAGCATCTTTCGTTCTGCAGT 294
DB 21859 CGGTCATCGGAGCTTCTGTGAAACTGTAGAGCATGAGTGGCTTTTGTAGCTCTGCTCG 21918
QY 295 GAATGCTCTCATGATCTGCGAGATAGTGGTATTGCTTCTAGTAACAGCTCGTCTTC 354
DB 21919 GAAAGTTTAGCAAGTCTTTTCGATAAGAT----- 21948
QY 355 TACTAGCAGATCTGCAGCAGTGGGACTCAACGACAGCGACCGACCTAGCGCTCTCTCCACC 414
DB 21949 -----TTCTCTGTTCTTTTCAGGGAACGACTCTCTCA 21981
QY 415 CAGTTTGATGATTAAGACTCAAGCGCAAAACAGCTTACGATCTATCTTTTACCTCAAC 474
DB 21982 AACTTTTGATGAAGCTAAGACGCAAGCAGAGAGTGCAGAACTCGCTGCGAGATGCGAC 22041
QY 475 ATCACTAGCTCACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACCTAATATAA 534
DB 22042 TACTTATGATCAGTTCAAGCCGCTTTTACAGCAGCTGCAAGATGCTGTGAAACAGATGGA 22101
QY 535 GGAATACAGCGGCTACTGATAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAATGC 594
DB 22102 GCAATTAGCTACTACTGATGAGAAAAAGCTACAGTTGCTACATGGAACGGCTCTTGA 22161
QY 595 CGATGCAGTTAAAGTTGGCGCGCAANTACAGATTAGCGAAATATGCTTCGGATAACCA 654
DB 22162 GCGCAGAAAGTAGTACGCTGGATACACATAACCAAGCTGGGTGCTATTCTTACAGAAACA 22221
QY 655 AGCGATTTCTTCACTCTTTTAGTAAACTGACITTCCTTGCACCTCTTACAGGCTGCTCTTCT 714
DB 22222 GAACTTCTTGAGGCAATAAGACGACCTCTGCTATGATCAGATTATGGAGCTGCGGG 22281
QY 715 CCAATCTGTAGCAAAACAATAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCC 774
DB 22282 ACAAGTAGAACCAATAAACAACCTGCTGAGGAGCTTAATTAACAGTTGAAGGA----- 22335
QY 775 AGTAGTCCAGGGAAGAGCGCTGCAATTGCTCAATCTTTTAGTTGATCAGACAGATGTAC 834
DB 22336 -----AGCTGGGGTTAGCTATCTCTGTGATAGATGACCTTTGAGAAGCAAAATTACACCTC 22389
QY 835 AGCGACACAGATAGAGAAAGATGAAATGCGATTAGGATGCAATATTTTGCAGGACAGAA 894

```

```
Db 22390 AGGAACCTCAGGTTACTGAATTAGCAGATGCTATATCGGAAGCTTATGCTCGGGGAAAA 22449
Qy 895 CGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATATAGCAACATAGATTTCAGC 954
Db 22450 CAGTACCGCGCTGTGGGCAAGCACAGGCAAAATAACAGCCCGCAATATAGAGCTTC 22509
Qy 955 TAAAGCAGCATCCCTACTGCTAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGA 1014
Db 22510 CAAACAAACTATTGCAATGCAAAAAGTCATAGAGACGCTCTTAAACTTGTCTCCAGA 22569
Qy 1015 CTCTCCAAATCTTCAAGAAGCGGAACAAATGGTAAATACAGGCTCAGAAAGATCTTAAAA 1074
Db 22570 TTCTCCGATACTCAAGCTGCTTGAAGAACAACAACAGCAGCAAAAGATATCTCTCAA 22629
Qy 1075 TATCAAACTCGCAGATGGTTCTGTAGTGTCCAAATCCAGGAAC-----TACAGT 1122
Db 22630 TGTGAACCTTAGTGGTGTAGTGTGCTATCGGTGGTCTCGGAGCTCCTGGTAGTGT 22689
Qy 1123 TGGAGGCTCCAAGCAACAGGAAGTAGTATGTTAGTATGCTGTTTCCATGCTGTTAGA 1182
Db 22690 GGGGACTTCTCAAAATCGCGTGTCTACCTTAGGGGAAGTTTCGGGTATCGATGTTATTGAC 22749
Qy 1183 TGATGCTGAAAATGAGACCGCTTCCATTTTGATGCTCTGGGTTTCGTCAGATGATTCACAT 1242
Db 22750 TGATGTTGATAATGAACCCGACGAGATCATTTGCAAGGTTTCAGAAATATGATCGATAA 22809
Qy 1243 GTTCAATACGGAATACTGATTTCTCAAGTGCACAGGAGCTCCAGCAACAGCTA- 1301
Db 22810 CTTCCATGATCAAAACTCTGATTTTACAGCGCCTTTTGAAGAGATTATGAATCAAGTAAC 22869
Qy 1302 -----GAGCAGCGAAAGCGCTGGAGATGACAGTGTGCTCAGCGCTGGCAGATGC 1353
Db 22870 CGACTTATCAACGAGATCAATCTCGAGATGCGGAAGCTACAGCAACCTACAAGAAAT 22929
Qy 1354 TCAGAAAGCTTTAGAGCGGCTTAGGTAAAGCTGGGCAACACAGGGGCATCTCAATGC 1413
Db 22930 ACAACAAACCATACAAGATGCCCTTCAAGGAGCTGCCGCTCAAGACGGCATGATCAATGC 22989
Qy 1414 TTTAGGACAGATCGCTTCTGCTGCTGTGTGAGCGCAGGAG 1454
Db 22990 TTTAGGAGCTATAACAACAGCAGCTTCAATTTCTACAGGAG 23030
```

```
RESULT 15
BD243971
LOCUS BD243971 1550 bp DNA linear PAT 17-JUL-2003
DEFINITION Chlamydia antigens and corresponding DNA fragments and uses
thereof.
ACCESSION BD243971
VERSION BD243971.1 GI:33053741
KEYWORDS JP 2002528081-A/1.
SOURCE Chlamydia pneumoniae
ORGANISM Chlamydia pneumoniae
REFERENCE 1 (bases 1 to 1550)
AUTHORS Murdin,A.D., Oomen,R.P. and Dunn,P.L.
TITLE Chlamydia antigens and corresponding DNA fragments and uses thereof
JOURNAL Patent: JP 2002528081-A 1 03-SEP-2002;
CONNAUGHT LABORATORIES LTD
OS Chlamydia pneumoniae
PN JP 2002528081-A/1
PD 03-SEP-2002
PF 28-OCT-1999 JP 2000578453
PR 28-OCT-1998 US 60/106037,20-SEP-1999 US 60/154658 PR
26-OCT-1999 US 09/427501
PI ANDREW DAVID MURDIN,RAYMOND PETER OOMEN,PAMELA LESLEY DUNN PC
C12N15/09,C12N15/09,A61K39/00,A61K39/118,A61K39/39, PC
A61P9/10.
PC A61P11/00,A61P11/06,C07K14/22,C07K14/295,C07K16/12,C12N1/15, PC
C12N1/19,
PC C12N1/21,C12N5/10,C12P21/02,C12Q1/68,G01N33/53,G01N33/566, PC
G01N33/569//
```

```
PC (C12N1/21,C12R1:01),(C12N15/09,C12R1:01),(C12P21/02,C12R1:01),
PC C12N15/00,
PC C12N15/00,C12N5/00,A61K37/02,(C12N15/00,C12R1:01) CC
Chlamydia antigens and corresponding DNA fragments and uses CC
thereof
FH Key Location/Qualifiers
FT CDS (101)..(1366).
FEATURES
source Location/Qualifiers
1..1550
/organism="Chlamydia pneumoniae"
/mol_type="genomic DNA"
/db_xref="taxon:83558"
```

ORIGIN

```
Query Match 11.7%; Score 170; DB 6; Length 1550;
Best Local Similarity 100.0%; Pred. No. 7.3e-26;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATAAATCTTTAAACACGCTCGCATTAATTATTAGTAGGAGCTTTTATTTTATTTT 60
Db 1381 ATAAATCTTTAAACACGCTCGCATTAATTATTAGTAGGAGCTTTTATTTTATTTT 1440
Qy 61 ATAAATCTTTAAACACGCTCGCATTAATTATTAGTAGGAGCTTTTATTTTATTTT 120
Db 1441 ATAAATCTTTAAACACGCTCGCATTAATTATTAGTAGGAGCTTTTATTTTATTTT 1500
Qy 121 AGTCTCTATAGAGAAACAGACGACACCTCCGCGAGATCTTTCTGCTC 170
Db 1501 AGTCTCTATAGAGAAACAGACGACACCTCCGCGAGATCTTTCTGCTC 1550
```

Search completed: March 24, 2004, 01:57:16
Job time : 5663.52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:19 ; Search time 3817.47 Seconds

(without alignments)
11389.568 Million cell updates/sec

Title: US-10-608-559-5

Perfect score: 1456

Sequence: 1 ataaaatttttaaacagg.....ctgttgtgagcaggaacta 1456

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estro:*

7: em_estpl:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	52.8	3.6	843	28	BZ643413	BZ643413. OGAOW57TM
C 2	52	3.6	1201	9	AL541034	AL541034. OGAOW57TM
C 3	50	3.4	1261	13	BX446742	BX446742. OGAOW57TM
C 4	49	3.4	909	13	BX453105	BX453105. OGAOW57TM

C 5	48.4	3.3	1191	14	CD500670	CD500670. CDR46-D01
C 6	48.2	3.3	533	13	EX353273	EX353273. EX353273
C 7	47.8	3.3	510	28	BH714761	BH714761. BOMGV1TR
C 8	47	3.2	808	29	CG815585	CG815585. SOYFPH83TV
C 9	46.8	3.2	1201	13	EX356412	EX356412. EX356412
C 10	46.6	3.2	1201	9	AL531906	AL531906. AL531906
C 11	46.6	3.2	1202	28	CC262481	CC262481. CH261-167
C 12	46	3.2	427	13	EX403499	EX403499. BX403499
C 13	46	3.2	1041	9	AL575932	AL575932. AL575932
C 14	46	3.2	1201	9	AL558113	AL558113. AL558113
C 15	45.8	3.1	642	13	EX396222	EX396222. BX396222
C 16	45.8	3.1	1104	13	EX396222	EX396222. BX396222
C 17	45.8	3.1	1201	9	AL561244	AL561244. AL561244
C 18	45.6	3.1	967	28	BH184682	BH184682. O2S_N_14-
C 19	45.6	3.1	967	29	CNS07PKF	AL621633. I3_end_01
C 20	45.4	3.1	528	12	BJ340195	BJ340195. BJ340195
C 21	45.4	3.1	724	12	BJ344100	BJ344100. BJ344100
C 22	45.4	3.1	819	28	BZ391755	BZ391755. EINH26TR
C 23	45.4	3.1	1201	9	AL524496	AL524496. AL524496
C 24	45.2	3.1	336	14	CB098017	CB098017. ku48b04.Y
C 25	45.2	3.1	1004	13	EX334499	EX334499. BX334499
C 26	45.2	3.1	1072	13	EX359378	EX359378. BX359378
C 27	45	3.1	1164	28	CC261365	CC261365. CH261-72H
C 28	44.8	3.1	1124	13	EX436282	EX436282. BX436282
C 29	44.8	3.1	1151	10	BB964820	BB964820. 601658475
C 30	44.8	3.1	1173	9	AL551058	AL551058. AL551058
C 31	44.8	3.1	1183	13	EX379850	EX379850. BX379850
C 32	44.6	3.1	259	14	CB722003	CB722003. jhm603F04
C 33	44.6	3.1	1537	12	BH195792	BH195792. 602754186
C 34	44.4	3.0	856	28	BH132709	BH132709. ENTNGS9TF
C 35	44.4	3.0	976	13	EX388320	EX388320. BX388320
C 36	44.4	3.0	1096	13	EX381337	EX381337. BX381337
C 37	44.4	3.0	1154	28	BZ636556	BZ636556. SP_Ba008
C 38	44.4	3.0	1201	9	AL548036	AL548036. AL548036
C 39	44.2	3.0	518	13	EX355479	EX355479. BX355479
C 40	44.2	3.0	591	9	AU286627	AU286627. AU286627
C 41	44.2	3.0	1201	9	AL531430	AL531430. AL531430
C 42	44.2	3.0	1201	13	EX458906	EX458906. BX458906
C 43	44	3.0	1174	9	AL546393	AL546393. AL546393
C 44	44	3.0	1201	13	EX334085	EX334085. BX334085
C 45	44	3.0	1201	13	EX364759	EX364759. BX364759

ALIGNMENTS

RESULT 1
BZ643413/c
LOCUS OGAOW57TM ZM 0.7_1.5 KB Zea mays genomic clone ZMMEM011J17,
DEFINITION genomic survey sequence.
ACCESSION BZ643413
VERSION BZ643413.1 GI:28104915
KEYWORDS GSS,
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 843)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGAOW57TC
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR

```

Class: sheared ends.
Location/Qualifiers
1. .843
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="W73"
/db_xref="taxon:4577"
/clone_lib="ZM00111117"
/notes="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 3.6%; Score 52.8; DB 28; Length 843;
Best Local Similarity 47.8%; Pred. No. 2.3; Mismatches 0; Gaps 0;
Matches 153; Conservative 0;

QY 286 TTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTCTAGTAACAG 345
DB 742 TACTGTACTACTGCTACTGCTGCTACTGCTACTGCTGCTACTGCTACTACTACTGCTAC 693
QY 346 CTGCTCTTCTACTAGCAGATCTGCAGACGFGGACTCAACAGACAGCGACCTAGGCC 405
DB 682 TACTGTACTACTACTGCTACTGCTGCTACTACTACTACTACTGCTACTGCTACTGCTGC 623
QY 406 TCTCTCCACCCAGCTTTGATGATTATAAGACTCAAGCGCAACAGCTTAACGATATCTT 465
DB 622 TACTGTCTACTGCTGCTACTGCTGCTACTGCTGCTACTGCTGCTGCTGCTACTGCTGC 563
QY 466 TACTCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTCTAC 525
DB 562 TGCTGTCTACTGCTACTGCTGCTACTGCTGCTGCTACTGCTGCTACTGCTACTGCTGC 503
QY 526 TAATATAAGAGTACAGCGCTACTGATGAGGAAACCGGAATCGTGCAGAGTGGGAAC 585
DB 502 TACTACTACTGCTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTACTGCTACTGCTAC 443
QY 586 TAAGATGCGCGAGCAGTTA 605
DB 442 TACTACTACTGCTGCTGCTA 423

RESULT 2
AL541034/c 1201 bp mRNA linear EST 12-MAY-2003
LOCUS AL541034 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005Y21
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION AL541034
VERSION AL541034.2 GI:30544815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12871720.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 951.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE005AALINPI&cluster=951.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue, Genoscope sequence ID : CS0DE005AALINPI.

FEATURES
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB011ZG07"
/tissue_type="NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
source

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005Y21"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 3.6%; Score 52; DB 9; Length 1201;
Best Local Similarity 42.0%; Pred. No. 3.1; Mismatches 0; Gaps 0;
Matches 63; Conservative 35;

QY 1 ATAAATCTTTAAACACAGCTCGCATTAAATATTAGTGAGAGCTTTTTTTTATTTTTT 60
DB 1160 AAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1101
QY 61 ATAAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 120
DB 1100 WAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1041
QY 121 AGGTCTCTATAGACGAAACAGACGACACC 150
DB 1040 VMMMMMMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1011

RESULT 3
BX446742/c 1201 bp mRNA linear EST 22-MAY-2003
LOCUS BX446742 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION CL0BB011ZG07 5-PRIME, mRNA sequence.
ACCESSION BX446742
VERSION BX446742.1 GI:31025761
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8247.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BB011ZG07RPI&cluster=8247.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue, Genoscope sequence ID : CL0BB011ZG07RPI.

FEATURES
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB011ZG07"
/tissue_type="NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
source

```

823 764

ORIGIN	Query	Match	Score	48.4;	DB	14;	Length	1191;
EST sequencing.								

Query Match 3.3%; Score 48.4; DB 14; Length 1191;
Best Local Similarity 48.5%; Pred. No. 17;
Matches 133; Conservative 0; Mismatches 141; Indels 0;


```

QY 675 GTAACTGACTCTTCGACCTCTTACAGGCTGCTCTTCTCCATCTGTAGCAACAATA 734
Db 858 GTGATATTTTGCAATAGCCGATAAAGTCTGAAAGACAGAGACTGAAGCAGTGCACG 799
QY 735 ACAAGCAGCTGAGCTCTTTAAAGAGATGCAAGATAACCCAGTAGTCCACAGGGAACGC 794
Db 798 CCACAGATATCAAGCCTTTTAAAGACGAGATGGAGAGAGATGTAAGAGAGCATATTC 739
QY 795 CTGCAATTCGTCATCTTTAGTGTATCAGACAGATGCTCAGCAGCACACATAGAGAAG 854
Db 738 CCACTATTTGACATCTTACAGATCAAAACCATCTGATAAGTGATTTAGAGAAAGAAAG 679
QY 855 ATGGAATGCGATTAGGATGTCATATTTTCAGACACAGAAGCGTAGTGGAGCTGTAGAA 914
Db 678 AGGAAAGAGATGGCGGGGAAAGGAAGAGAGAGAAATCAATAGAGAAAGAAAGTA 619
QY 915 ATGCTAAATCTAATAACAGTATTAAGCAACATAGA 948
Db 618 TTGTAAGAGTGATACGTCTGCTACCAATAAAGA 585

RESULT 6
BX353273/c
LOCUS
DEFINITION
BX353273 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC003YH11 5-PRIME, mRNA sequence.
ACCESSION
BX353273
VERSION
BX353273.1 GI:30377779
KEYWORDS
Esi.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 533)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC003CD06QP1.
FEATURES
Location/Qualifiers
i. .533
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC003YH11"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 3.3%; Score 48.2; DB 13; Length 533;
Best Local Similarity 41.4%; Pred. No. 21;
Matches 48; Conservative 35; Mismatches 33; Indels 0; Gaps 0;
QY 1 ATAAATCTTTAAACACAGCTCGCATTAATTAATAGTGAGAGCTTTTTTTTATTTTTT 60
Db 431 AAAAAATATWWWWTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTW 372
QY 61 ATATAAACAATAAGATTTTATTTATTTTTCAGTTTTTATGTTAATCTAATG 116
Db 371 WWWWWWWWWWWHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 316

RESULT 7

```

```

BH714761
LOCUS
DEFINITION
BOMGV71TR BO 2.3 KB Brassica oleracea genomic clone BOMGV71,
genomic survey sequence.
ACCESSION
BH714761
VERSION
BH714761.1 GI:19808752
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 510)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOMGV71TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
i. .510
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T0100DH3"
/db_xref="taxon:3712"
/clone="BOMGV71"
/clone_lib="BO_2.3 KB"
/notes="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN
Query Match 3.3%; Score 47.8; DB 28; Length 510;
Best Local Similarity 63.5%; Pred. No. 25;
Matches 73; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 5 AATCTTTTAAAAACAGCTCGCATTAATTAATAGTGAGAGCTTTTTTTTATTTTATAA 64
Db 388 ATAGTTAAAACTGTACTACTAGTTTTTGGAGTTTTTTTTTTTTTTTGTCTAA 447
QY 65 TAAACTAAAGATTTTATTAATTTTGGTTTTTATGTTTAACTTCATTGTC 119
Db 448 TAGCTTTTCGGAGTTTCAAGTTTGATGAGTTGTTTGGAGCTTCTAATGTC 502

RESULT 8
CG815585/c
LOCUS
DEFINITION
SOYFH83TV largeInsertSoybeanGenLib Glycine max genomic clone
H46C04:SEQN22, genomic survey sequence.
ACCESSION
CG815585
VERSION
CG815585.1 GI:38269059
KEYWORDS
GSS.
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 808)
Shultz, J., Meksem, K., Shetty, J., Town, C.D., Koo, H., Potter, J.,
Wakefield, K., Zhang, H., Wu, C. and Lightfoot, D.A.
End sequencing of BACs comprising a provisional minimal tiling path
from a fingerprint physical map of soybean (Glycine max) cultivar
Forrest
Unpublished (2003)
Other GSSs: SOYFH83TH

```

Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
The Center of Excellence in Soybean Research, Teaching and
Outreach, Southern Illinois University at Carbondale and Plant
Genomics, The Institute for Genomic Research
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 618 453 1797
Email: ga4082@siu.edu, jshultz@siu.edu, cdtown@tigr.org (URL:
http://bioinformatics.siu.edu)
Clones approximating a minimum tiling path were re-arrayed from the
library master plates prior to sequencing.
For purposes of clone identification each clone name is a
concatenation of the original clone location and its new location
in the re-arrayed sequencing plates.
Seq primer: GTAAACGACCTACTATAGGC
Class: BAC ends.

FEATURES

Location/Qualifiers
1..808
/organism="Glycine max"
/mol_type="genomic DNA"
/culivar="Forrest"
/db_xref="taxon:3847"
/clone="H46C04:SEQ322"
/note="Organ: Leaves; Vector: pCLD04541 (pBLOBAC11 EcorI
clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
cv. Forrest seeds were grown in greenhouse for fourteen
days. Nuclei were isolated and embedded in agarose,
restriction digested with Hind III BamHI or EcorI, large
size DNA fragments were ligated in vector V41 (pCLD04541)
and electro transformed in DH10a cells. About 90,000
clones from BAC libraries were fingerprinted with HindIII
and Hae III. Version 2 (automatic build) Contigs were
built from 78,001 fingerprints. Contigs were manually
examined to find the best non redundant tile path through
the contigs representing about 13,000 clones. The clones
were end sequenced."

ORIGIN

Query Match 3.2%; Score 47; DB 29; Length 808;
Best Local Similarity 64.0%; Pred. NO. 34;
Matches 71; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 3 AAAATCTTTAAACAGCTCGCATTAATTAGTGAGAGCTTTTATTTTATTTAT 62
Db 625 AAATCTTCAAGAAAAAGCTCGATTGATTTTCTTGATTTTATTTATTTT 566
QY 63 AATAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 113
Db 565 CAAGATATTTTGATTTATTTATTTATTTTATTTTATTTTATTTTATTT 515

RESULT 9
BX356412/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BX356412
BX356412 Homo sapiens PLACENTA COT 25-NORMALIZED
clone CSODI011YJ12 5-PRIME, mRNA sequence.
BX356412
BX356412.1 GI:30382071
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 1201)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI011DE06Qp1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI011YJ12"
/tissue type="PLACENTA COT 25-NORMALIZED"
/note="lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double strand cDNA was
digested with Not I and cloned into the Not I and Ecor V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.2%; Score 46.8; DB 13; Length 1201;
Best Local Similarity 27.2%; Pred. NO. 35;
Matches 66; Conservative 75; Mismatches 102; Indels 0; Gaps 0;
QY 55 TTTTATAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTT 114
Db 914 TTTTATAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTT 855
QY 115 TGTCCAGGTCCTATAGACGAAACAGACGACACCTCCCGCAGATCTTTCTGCTCAAG 174
Db 854 AAATTTTAAATAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 795
QY 175 ATTGAGGCGAGTCAGCAATAAGAGTCGGAAGCTCAAGAATAGCAGTGCGGAAGC 234
Db 794 KKK 735
QY 235 TAAGCTTAAGAACTAAGACCGATTTCTAGACGATGGAGCATCTTGGCTTGCAGT 294
Db 734 ADAAACAATAAAACCKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 675
QY 295 GAA 297
Db 674 AAA 672

RESULT 10
AL531906 1201 bp mRNA linear EST 23-MAY-2003
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
AL531906
AL531906 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CSODM003YL08 3-PRIME, mRNA sequence.
AL531906
EST.
GI:31069738
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 1201)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODM003DF04NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM003YL08"

FEATURES

Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM003YL08"

```

/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match          3.2%; Score 46.6; DB 9; Length 1201;
Best Local Similarity 48.2%; Pred. No. 38;
Matches 67; Conservative 19; Mismatches 53; Indels 0; Gaps 0;

QY 28 TAATTATTAGTGAGAGCTTTTATTTTATTTTATTAATAAACTAAAGATTTTATTAT 87
Db 670 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 729
QY 88 TTTTTCAGTTTATGTTTATCTCTATTGTCAGGTCCTATAGACGAAACAGACGAC 147
Db 730 TTTTATTTTATTAAGATTAATAAAAGAAACAGCTTTATATATATATATATATATAT 789
QY 148 ACCTCCGAGATCTTCT 166
Db 790 ATWTAATAATACATTTTAT 808

RESULT 11
LOCUS CC262481
DEFINITION CH261-167M9 Sp6.1 CH261 Gallus gallus genomic clone CH261-167M9,
genomic survey sequence.
ACCESSION CC262481
VERSION CC262481.1 GI:30607397
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 1202)
AUTHORS Archosaura; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Phasianidae; Gallus.
TITLE Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
JOURNAL Gallus gallus BAC End Reads
COMMENT Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 30
High quality sequence stop: 105.
FEATURES
Location/Qualifiers
1..1202
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-167M9"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/notes="Vector: pFARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN
Query Match          3.2%; Score 46.6; DB 28; Length 1202;
Best Local Similarity 58.2%; Pred. No. 38;
Matches 82; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match          3.2%; Score 46; DB 13; Length 427;
Best Local Similarity 31.5%; Pred. No. 59;
Matches 34; Conservative 47; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATAAATCTTTTAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTATTTTATTTT 60
Db 1004 ATAAATCTTTTAAATTAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1063
QY 61 ATAAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 120
Db 1064 ATAAATAAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1123
QY 121 AGGTCCTATAGACGAAACAGA 141
Db 1124 ATTATATTAAATATAAATA 1144

RESULT 12
LOCUS BX403499
DEFINITION BX403499 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA002ZH05
3-PRIME, mRNA sequence.
ACCESSION BX403499
VERSION BX403499.1 GI:30635017
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 427)
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA002ZH05FP1.
FEATURES
Location/Qualifiers
1..427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA002ZH05"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match          3.2%; Score 46; DB 13; Length 427;
Best Local Similarity 31.5%; Pred. No. 59;
Matches 34; Conservative 47; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATAAATCTTTTAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTATTTTATTTT 60
Db 64 WWWWWWWWWWWTWWTWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 123
QY 61 ATAAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 108
Db 124 WTWWTAAATWAATATWATWATWATWATWATWATWATWATWATWATWATWATWATW 171

RESULT 13
LOCUS AL575932/c
DEFINITION AL575932 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1067YN05 3-PRIME, mRNA sequence.
ACCESSION AL575932
VERSION AL575932.2 GI:31314228
KEYWORDS EST.

```


/note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dT primed from Suq of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match	3.1%;	Score 45.8;	DB 13;	Length 642;
Best Local Similarity	50.7%;	Pred. No. 61;		
Matches 110;	Conservative	0;	Mismatches 107;	Indels 0;
			Gaps	0;
QY	45	TTTTTTTTTTTATATATAAAACCTAAAGATTTTATTTTGTAGTTTTTATGG	104	
Db	640			
QY	105	TTAATCCTATTGTCAGGTCCTATAGACGAAACAGAACCCACACCTCCCGCAGATCTTT	164	
Db	580	GGTTACCAATTTTGTGCTGTTTCGAGATCTACAGTGAACACAAATCCGGTAGTCTGA	521	
QY	165	CTGCTCAAGGATTGGAGCGAGTGCGACAAATAAGAGTGGGAGCTCAAAGATAGCAG	224	
Db	520	CCATTCTCGGCTTAGCTTAAGTTTAAAGCGATACAGCCGAGGAACATGTAAACATACCTC	461	
QY	225	GTGCGAAGCTAAGCCTAAGAACTCTAAGACCGATTC	261	
Db	460	AGTCCTAACCTTTATGGAATTATTCATCCACCGATAC	424	

Search completed: March 24, 2004, 05:56:12
Job time : 3820.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:04 ; Search time 830.202 Seconds
(without alignments)
11032.393 Million cell updates/sec

Title: US-10-608-559-1

Perfect score: 2156

Sequence: 1 ataaatttttaaacagg.....gatcttcaataataaaaga 2156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002as.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2156	100.0	2156	3 AAD02063	Aad02063 Chlamydia
2	2143.4	99.4	110000	2 AAX91990_08	Continuation (9 of
3	1956	90.7	1956	4 AAS57031	Aas57031 C. pneumo
4	1956	90.7	1956	6 ABL92612	Ab192612 Chlamydia
5	1956	90.7	1956	6 ABL91190	Ab191190 Chlamydia
6	1956	90.7	1956	9 ADD42825	Add42825 Chlamydia
7	1848.8	85.8	1852	3 AAD02064	Aad02064 5'-trunca
8	1465.8	68.0	2238	3 AAD02066	Aad02066 C. pneumo
9	1455	67.5	1456	3 AAD02065	Aad02065 3'-trunca
10	187.6	8.7	110000	2 AAZ01425_07	Continuation (8 of
11	187.6	8.7	110000	2 AAZ01425_08	Continuation (9 of
12	187.2	8.7	1537	3 AAB63292	Aab63292 C. tracho
13	187.2	8.7	1537	4 AAB56196	Ab192425 Chlamydia
14	187.2	8.7	1537	6 ABL92425	Ab192425 Chlamydia
15	186	8.6	1944	9 ADD43867	Add43867 Chlamydia
16	185.4	8.6	1941	6 ABL92619	Ab192619 Chlamydia
17	184	8.5	1171	4 AAS56981	Aas56981 C. tracho
18	184	8.5	1171	9 ADD42775	Add42775 Chlamydia
19	184	8.5	1834	4 AAS57006	Aas57006 C. tracho
20	184	8.5	1834	9 ADD42800	Add42800 Chlamydia
21	183	8.5	1983	4 AAS56996	Aas56996 C. tracho
22	183	8.5	1983	9 ADD42790	Add42790 Chlamydia
23	170	7.9	1550	3 AAB28411	Aab28411 Chlamydia

24	148.4	6.9	150	2	AAV16207	Aav16207 Part of t
25	55.4	2.6	2000	7	ADA71938	Ada71938 Rice gene
26	51.6	2.4	2205	8	ADA28999	Ada28999 DNA encod
27	49	2.3	896	2	AAQ68903	Aaq68903 PspA prot
28	49	2.3	946	2	AAQ33124	Aax33124 Streptoco
29	49	2.3	957	2	AAT61726	Aat61726 Streptoco
30	49	2.3	1860	7	ACA49648	Aca49648 Prokaryot
31	49	2.3	1990	7	AAT61725	Ast61725 Streptoco
32	49	2.3	1990	7	ABX95373	Abx95373 S. pneumo
33	49	2.3	2085	2	AAQ78131	Aaq78131 Pneumococ
34	49	2.3	2085	2	AAT08979	Aat08979 DNA encod
35	49	2.3	2085	2	AAT07178	Aat07178 Pneumococ
36	49	2.3	2085	2	AAV33264	Aav33264 Pneumococ
37	49	2.3	2085	2	AAV39470	Aav39470 Streptoco
38	49	2.3	2085	2	AZ25063	Aaz25063 Streptoco
39	49	2.3	2085	2	AAV84069	Aav84069 DNA encod
40	47.4	2.2	2085	2	AAQ02012	Aax02012 S. pneumo
41	47.4	2.2	2085	7	ABX95374	Abx95374 S. pneumo
42	47.4	2.2	2085	2	AAQ28674	Aaq28674 PspA gene
43	47.2	2.2	110000	6	ABA90521_07	Continuation (8 of
44	47	2.2	9373	6	ABL33291	Ab133291 Human imm
45	47	2.2	9373	6	ABK31317	Abk31317 Signal tr

ALIGNMENTS

RESULT 1

AAD02063 AAD02063 standard; DNA; 2156 BP.
XX

AC AAD02063;
XX

DT 15-SEP-2003 (revised)
DT 26-MAR-2001 (first entry)
XX

DE Chlamydia pneumoniae 76 kDa full-length protein encoding DNA.
XX

XX 76 kDa protein; bactericidal; diagnosis; prevention; treatment;
KW pneumonia; upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; ds.
XX

OS Chlamydia pneumoniae.
XX

EH Key Location/Qualifiers
FT CDS 101..2056
FT FT /*tag= a
FT FT /product= "Chlamydia pneumoniae 76kDa protein"

XX WO200066739-A2.
XX

XX 09-NOV-2000.
XX

XX 03-MAY-2000; 2000WO-CA000511.
XX

XX 03-MAY-1999; 99US-0132270P.
XX

XX 30-JUN-1999; 99US-0141276P.
XX

XX (AVET) AVENTIS PASTEUR LTD.
XX

XX Murdin AD, Oomen RP, Wang J, Dunn P;
XX

XX WPI; 2000-687542/67.
XX

XX P-PSDB; AAY71954.
XX

XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
XX for vaccinating against Chlamydia infections.
XX

XX Claim 2a; Fig 1; 112pp; English.
XX

XX The present sequence is a DNA coding for Chlamydia pneumoniae 76 kDa
XX protein. C. pneumoniae 76 kDa protein is used in the diagnosis,
XX

Original

CC prevention and treatment of *C. pneumoniae* infections (e.g. pneumonia,
CC upper respiratory tract disease, bronchitis, sinusitis and acute
CC respiratory disease such as cough, sore throat, hoarseness, fever; and
CC abnormal chest sounds on auscultation). *C. pneumoniae* sequence is also
CC used as vaccines for immunising humans against diseases caused by *C.*
CC *pneumoniae*. (Updated on 15-SEP-2003 to standardise OS field)

Sequence 2156 BP; 672 A; 461 C; 471 G; 552 T; 0 U; 0 Other;
XX

Query Match 100.0%; Score 2156; DB 3; Length 2156;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 0; Indels 0;

QY	1	ATATAAATCTTTAAAAACAGGCTCGAATTAATATTAGTGAGAGCTTTTATTTTATTTTATTTT	60
DB	1	ATATAAATCTTTAAAAACAGGCTCGAATTAATTAATTAGTGAGAGCTTTTATTTTATTTTATTTT	60
QY	61	ATAATAAAACTAAAAAGATTTTATTATTTTGTAGTTTTTATGCTTAATCCTATGCTGCC	120
DB	61	ATAATAAAACTAAAGATTTTATTATTTTGTAGTTTTTATGCTTAATCCTATGCTGCC	120
QY	121	AGTTCCTATAGACGAAACAGAACGCAACCTCCCGCGAGATCTTCTGCTCAAGGATGGA	180
DB	121	AGTTCCTATAGACGAAACAGAACGCAACCTCCCGCGAGATCTTCTGCTCAAGGATGGA	180
QY	181	GCGGAGTGCAGCAAAATAAGATGCGGAAGCTCAAGAAGATAGCAGGTGCGGAAGCTAAGCC	240
DB	181	GCGGAGTGCAGCAAAATAAGATGCGGAAGCTCAAGAAGATAGCAGGTGCGGAAGCTAAGCC	240
QY	241	TAAAGAAATCTAAGACCGATTCGTAGACGCGATGGAGCATCTTCGTTTCTGCAGTGAATGC	300
DB	241	TAAAGAAATCTAAGACCGATTCGTAGACGCGATGGAGCATCTTCGTTTCTGCAGTGAATGC	300
QY	301	TCTCATGAGTCTGCAGATAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG	360
DB	301	TCTCATGAGTCTGCAGATAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG	360
QY	361	CAGATCTGCAGACGTTGGACTCAACGACAGCGACCGCACCTACGCTCTCCACCCACGTT	420
DB	361	CAGATCTGCAGACGTTGGACTCAACGACAGCGACCGCACCTACGCTCTCCACCCACGTT	420
QY	421	TGATGATTTAAGACTCAAGCGCAAAACAGCTTAGATACATCTTTACCTCAACATCATCT	480
DB	421	TGATGATTTAAGACTCAAGCGCAAAACAGCTTAGATACATCTTTACCTCAACATCATCT	480
QY	481	AGCTGACATACAGGCTGCTTTGGTGAGCGCTCCAGGATGCTGTCACATAATAAAGGATAC	540
DB	481	AGCTGACATACAGGCTGCTTTGGTGAGCGCTCCAGGATGCTGTCACATAATAAAGGATAC	540
QY	541	ACGGGTTACTGATGAGGAAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGC	600
DB	541	ACGGGTTACTGATGAGGAAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGC	600
QY	601	AGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATPAACCAACGAT	660
DB	601	AGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATPAACCAACGAT	660
QY	661	TCTTGACTCTTTAGTGAACTGACTTCCTTCGAGCTCTTAAGAGATGCAAGTAAACCCAGTAT	720
DB	661	TCTTGACTCTTTAGTGAACTGACTTCCTTCGAGCTCTTAAGAGATGCAAGTAAACCCAGTAT	720
QY	721	TGTAGCAACAATACAAAGCGCTGAGCTCTTTAAAGAGATGCAAGTAAACCCAGTAT	780
DB	721	TGTAGCAACAATACAAAGCGCTGAGCTCTTTAAAGAGATGCAAGTAAACCCAGTAT	780
QY	781	CCGAGGAAACCGCTGCAATTTGCTCAATCTTTTAGTTGATCAGACAGATGCTACAGCGAC	840
DB	781	CCGAGGAAACCGCTGCAATTTGCTCAATCTTTTAGTTGATCAGACAGATGCTACAGCGAC	840
QY	841	ACAGATAGAGAAAGATGGAATGCGAATAGGGATGCAATATTTTGCAGGACAGAAACGCTAG	900
DB	841	ACAGATAGAGAAAGATGGAATGCGAATAGGGATGCAATATTTTGCAGGACAGAAACGCTAG	900

Db	1981	CTTTGAAAGCACTCTCTGTTTATTACAGAGTGTCTGCTCAATATCGGCTCTCTATATTC	2040
Qy	2041	TGGTTATCTCCAAATACACACCTAAGTGTTCGTTTGGAGAGATTATTATGCTTTGGT	2100
Db	2041	TGGTTATCTCCAAATACACACCTAAGTGTTCGTTTGGAGAGATTATTATGCTTTGGT	2100
Qy	2101	AAGGCTTTGTTGAGGCTTTACCAACACACTAGACGATCTTCAATAATAAAGA	2156
Db	2101	AAGGCTTTGTTGAGGCTTTACCAACACACTAGACGATCTTCAATAATAAAGA	2156
RESULT 2			
AAx91990_08			
Continuation (9 of 13) of AAx91990 from base 800001 (Nucleotide sequence of the complete			
WP Sequence split into 13 fragments LOCUS AAx91990 Accession AAx91990			
WP	Fragment Name	Begin	End
WP	AAx91990_00	1	110000
WP	AAx91990_01	100001	210000
WP	AAx91990_02	200001	310000
WP	AAx91990_03	300001	410000
WP	AAx91990_04	400001	510000
WP	AAx91990_05	500001	610000
WP	AAx91990_06	600001	710000
WP	AAx91990_07	700001	810000
WP	AAx91990_08	800001	910000
WP	AAx91990_09	900001	1010000
WP	AAx91990_10	1000001	1110000
WP	AAx91990_11	1100001	1210000
WP	AAx91990_12	1200001	12300025
Query Match			
Best Local Similarity 99.4%; Score 2143.4; DB 2; Length 110000;			
Matches 2155; Conservative 99.9%; Pred. No. 0; Mismatches 1; Indels 1; Gaps 1;			
Qy	1	ATAAAATCTTAAACACAGGCTCGCATTAATATTAGTGAGAGCTTTTATTTTATTTT	60
Db	28698	ATAAAATCTTAAACACAGGCTCGCATTAATATTAGTGAGAGCTTTTATTTTATTTT	28639
Qy	61	ATAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	120
Db	28638	ATAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	28579
Qy	121	AGGTCCTATAGACGAAACAGAACCGACACCTCCCGCAGATCTTCTGCTCAAGGATTGGA	180
Db	28578	AGGTCCTATAGACGAAACAGAACCGACACCTCCCGCAGATCTTCTGCTCAAGGATTGGA	28519
Qy	181	GGCAGTGCAGCAATAAGAGTGGGAGCTCAAGAAATAGCAGTGGGAGCTTAAGCC	240
Db	28518	GGCAGTGCAGCAATAAGAGTGGGAGCTCAAGAAATAGCAGTGGGAGCTTAAGCC	28459
Qy	241	TAAAGATCTAAGACCGATTCGTAGACGATGGAGCATCTTGCGTCTGCAAGTGAATGC	300
Db	28458	TAAAGATCTAAGACCGATTCGTAGACGATGGAGCATCTTGCGTCTGCAAGTGAATGC	28399
Qy	301	TCTCATGAGTCTGCAGATAAGCTGGGATTTGCTTCTAGTAAACAGCTGCTTCTACTAG	360
Db	28398	TCTCATGAGTCTGCAGATAAGCTGGGATTTGCTTCTAGTAAACAGCTGCTTCTACTAG	28339
Qy	361	CAGATCTGCAGACGTGAGCTCAACGACGACGACGACCTACGCTCTCCACCCAGTT	420
Db	28338	CAGATCTGCAGACGTGAGCTCAACGACGACGACGACCTACGCTCTCCACCCAGTT	28279
Qy	421	TGATGATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT	480
Db	28278	TGATGATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT	28219
Qy	481	AGCTGACATACAGGCTGCTTTGGTGGAGCTTCAGGATGCTGCTCACTAATAATAAGGATAC	540
Db	28218	AGCTGACATACAGGCTGCTTTGGTGGAGCTTCAGGATGCTGCTCACTAATAATAAGGATAC	28159
Qy	541	AGCGGCTACTGATGAGGAAACCGCAATGCTGCGGAGTGGGAACCTAAGAAATGCCGATGC	600

Db	28158	AGCGGCTACTGATGAGGAAACCGCAATCGCTGCGGAGTGGAAACTAAGAAATGCCGATGC	28099
Qy	601	AGTTAAAGTTGCGCGCAAAATTACAGAAATTACGAAATATATCTTCGGATAACCAAGCGAT	660
Db	28098	AGTTAAAGTTGCGCGCAAAATTACAGAAATTACGAAATATATCTTCGGATAACCAAGCGAT	28039
Qy	661	TCATTGACCTTTAGTAAACTGACTCTCTTCGACCTCTTAAGGCTGCTCTTCTCCAATC	720
Db	28038	TCATTGACCTTTAGTAAACTGACTCTCTTCGACCTCTTAAGGCTGCTCTTCTCCAATC	27979
Qy	721	TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	780
Db	27978	TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	27919
Qy	781	CCCAGGAAACCGCTGCAATGCTCAATCTTAAAGAGATGCAAGATAACCCAGTAGT	840
Db	27918	CCCAGGAAACCGCTGCAATGCTCAATCTTAAAGAGATGCAAGATAACCCAGTAGT	27859
Qy	841	ACAGATAGAGAAAGATGGAATGCGATTTAGGATGCAATTTTTCAGGACAGACGCTAG	900
Db	27858	ACAGATAGAGAAAGATGGAATGCGATTTAGGATGCAATTTTTCAGGACAGACGCTAG	27799
Qy	901	TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGCAACATAGATTGCTTAAAGC	960
Db	27798	TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGCAACATAGATTGCTTAAAGC	27739
Qy	961	AGCAATCGTCTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC	1020
Db	27738	AGCAATCGTCTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC	27679
Qy	1021	AATTCTTCAAGAGCGGAACAAATGCTAAATACAGCTGAGAAAGATCTTAAATAATCAA	1080
Db	27678	AATTCTTCAAGAGCGGAACAAATGCTAAATACAGCTGAGAAAGATCTTAAATAATCAA	27619
Qy	1081	ACCTGCAAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACA	1140
Db	27618	ACCTGCAAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACA	27559
Qy	1141	AGGAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTAGATGATGCTGAAAATGAGAC	1200
Db	27558	AGGAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTAGATGATGCTGAAAATGAGAC	27499
Qy	1201	CGCTTCCATTTTGATGTTCTGTTTTCGATGATTTTCAATGATTTCAATCGGAAATCC	1260
Db	27498	CGCTTCCATTTTGATGTTCTGTTTTCGATGATTTTCAATGATTTCAATCGGAAATCC	27439
Qy	1261	TGATTCTCAAGCTGCCCAACAGGAGCTCCGAGCAACAGCTAGACGCAAGAGCCGCTGG	1320
Db	27438	TGATTCTCAAGCTGCCCAACAGGAGCTCCGAGCAACAGCTAGACGCAAGAGCCGCTGG	27379
Qy	1321	AGATGACAGTCTGCTGCGAGCTGCGAGATGCTCAGAAAAGCTTTAGAGCGGCTTAGG	1380
Db	27378	AGATGACAGTCTGCTGCGAGCTGCGAGATGCTCAGAAAAGCTTTAGAGCGGCTTAGG	27319
Qy	1381	TAAAGCTGGGCAACAAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTG	1439
Db	27318	TAAAGCTGGGCAACAAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTG	27259
Qy	1440	TTGTGAGCGCAGGAGTTCTCCCGCTGCGACCAAGTTCTATAGGGTCACTGTAAAAACAGC	1499
Db	27258	TTGTGAGCGCAGGAGTTCTCCCGCTGCGACCAAGTTCTATAGGGTCACTGTAAAAACAGC	27199
Qy	1500	TTTACAGAGCTCAAAATCTACAGGTTCTGATTTATAAACACAGATATCAGCAGGTTATG	1559
Db	27198	TTTACAGAGCTCAAAATCTACAGGTTCTGATTTATAAACACAGATATCAGCAGGTTATG	27139
Qy	1560	ATGCTTACAATCCATCAATGATGCTTATGTTAGGCGACGAAATGATGCCACTCGTGATG	1619
Db	27138	ATGCTTACAATCCATCAATGATGCTTATGTTAGGCGACGAAATGATGCCACTCGTGATG	27079
Qy	1620	TGATTAACAATGTAAGTACCCCGCTCTCACAGATCCGTTCCCTAGACGACGAAACGAAAG	1679
Db	27078	TGATTAACAATGTAAGTACCCCGCTCTCACAGATCCGTTCCCTAGACGACGAAACGAAAG	27019

QY 1680 CTCGAGCACCAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGCAATAGCAGAA 1739
Db 27018 CTCGAGCACCAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGCAATAGCAGAA 26959
QY 1740 CTCCTCGAGATGCTATAGTCAAGTTTTCGGCACTACAATCTGTAATGCGATCATCCAGT 1799
Db 26958 CTCCTCGAGATGCTATAGTCAAGTTTTCGGCACTACAATCTGTAATGCGATCATCCAGT 26899
QY 1800 CGAATCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGAGTGAACAAAGC 1859
Db 26898 CGAATCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGAGTGAACAAAGC 26839
QY 1860 CTCACAGTTTGGCTATCTTATGTCACACTTCTTAATGACTCTACACAGAGTTTCATAG 1919
Db 26838 CTCACAGTTTGGCTATCTTATGTCACACTTCTTAATGACTCTACACAGAGTTTCATAG 26779
QY 1920 CTAAATGAAAGTTTGGTTGCTGAAGGATCTAGGACAGCTGAATATAAAGCACATTT 1979
Db 26778 CTAAATGAAAGTTTGGTTGCTGAAGGATCTAGGACAGCTGAATATAAAGCACATTT 26719
QY 1980 CCTTTGAAAGCACTCTCTTGTATTTATTCAGCAGTCTGCTCAATATCGGCTCTCTATAT 2039
Db 26718 CCTTTGAAAGCACTCTCTTGTATTTATTCAGCAGTCTGCTCAATATCGGCTCTCTATAT 26659
QY 2040 CTGGTTATCTCCAATAAACACACCTAAGTGTTCGTTTGGAGAGATTTATTTGTTTGG 2099
Db 26658 CTGGTTATCTCCAATAAACACACCTAAGTGTTCGTTTGGAGAGATTTATTTGTTTGG 26599
QY 2100 TAAAGCCTTTGTTAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 2156
Db 26598 TAAAGCCTTTGTTAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 26542

RESULT 3
ID AAS57031 standard; DNA; 1956 BP.
AC AAS57031;
XX
XX
DT 11-SEP-2003 (revised)
DT 16-JAN-2002 (first entry)
XX
XX
DE C. pneumoniae DNA encoding the CT622 homologue CPn0728.
XX
KW Chlamydia; ds; sexually transmitted disease; PID; antibacterial;
KW pelvic inflammatory disease; antigen; trachoma; gynecological;
KW acute respiratory tract infection; atherosclerosis; male infertility;
KW coronary heart disease.
XX
OS Chlamydophila pneumoniae.
XX
PN WC940181329-A2.
PD 01-NOV-2001.
XX
PF 23-APR-2001; 2001WO-US013081.
XX
PR 21-APR-2000; 2000US-0198853P.
PR 20-JUL-2000; 2000US-0219752P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bhatia A, Probst P, Stromberg EJ;
XX
DR WPI; 2001-616771/71.
DR P-PSDB; AAU88899.
XX
PT New polynucleotide for treating Chlamydia infections encodes a
PT polynucleotides containing an immunogenic portion of a Chlamydia antigen.
PS Disclosure; Page 143; 208pp; English.
XX

The invention relates to isolated polynucleotide encoding at least a partial Chlamydia protein which is an antigenic fragment, or the complements, fragments, homologues and variants, and antibodies raised against the antigenic proteins (or fragments). The nucleic acids, proteins and antibodies are used to diagnose and treat Chlamydia infections (e.g. a sexually transmitted disease, pelvic inflammatory disease (PID), acute respiratory tract infection, trachoma, atherosclerosis and coronary heart disease) in a patient, and in the treatment of male infertility. The compounds of the invention are also useful for detecting the presence of Chlamydia in a patient, and stimulating and/or expanding T cells specific for a Chlamydia protein. The present sequence encodes a Chlamydia antigen. (Updated on 11-SEP-2003 to standardise OS field)

QY 101 ATGGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 160
Db 101 ATGGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 60
QY 161 CTTTCTGCTCAAGGATTCGAGCGAGTGCAGCAATAAGAGTGCAGGAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTCGAGCGAGTGCAGCAATAAGAGTGCAGGAGCTCAAGAATA 120
QY 221 GCAGTGGCGAAGCTAAGCCCTAAAGAAATCTAAGACCGATTTCTAGAGCGATGGAGCATC 280
Db 121 GCAGTGGCGAAGCTAAGCCCTAAGAAATCTAAGACCGATTTCTAGAGCGATGGAGCATC 180
QY 281 TTGGTCTGCAAGTGAATGCTCTCATGAGTCTGCGAGATAAGTGGGTATTTGCTTTAGT 340
Db 181 TTGGTCTGCAAGTGAATGCTCTCATGAGTCTGCGAGATAAGTGGGTATTTGCTTTAGT 240
QY 341 AACAGCTGCTCTTCTACTAGCAGATCTGCAGCTGGACTCAACGACGACGACCGCCT 400
Db 241 AACAGCTGCTCTTCTACTAGCAGATCTGCAGCTGGACTCAACGACGACGACCGCCT 300
QY 401 ACGCTCTCTCCACCCAGCTTTGATTAAGACTCAAGCGCAAAACAGCTTTAGCATCT 460
Db 301 ACGCTCTCTCCACCCAGCTTTGATTAAGACTCAAGCGCAAAACAGCTTTAGCATCT 360
QY 461 ATCTTTACCTCAACATCACTAGCTGACATACAGCTGCTTTGGTGGAGCTCCAGATGCT 520
Db 361 ATCTTTACCTCAACATCACTAGCTGACATACAGCTGCTTTGGTGGAGCTCCAGATGCT 420
QY 521 GTCCTAATAATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGGAGTGG 580
Db 421 GTCCTAATAATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGGAGTGG 480
QY 581 GAAACTAAGATGCGATGAGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 640
Db 481 GAAACTAAGATGCGATGAGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 540
QY 641 GCTTCGGATACCAAGCGATTTGACTCTTTAGTTAACTGACTTCTCTTCGACTCTTA 700
Db 541 GCTTCGGATACCAAGCGATTTGACTCTTTAGTTAACTGACTTCTCTTCGACTCTTA 600
QY 701 CAGGCTGCTCTTCTCCAACTGTAGCAAAACAAATAAAGCAGCTGAGCTTTCTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAACTGTAGCAAAACAAATAAAGCAGCTGAGCTTTCTAAAGAG 660
QY 761 ATGCAAGATTAACCCAGTAGTCCCGAGGAAACCGCTGCAATTCCTAATCTTTAGTTGAT 820
Db 661 ATGCAAGATTAACCCAGTAGTCCCGAGGAAACCGCTGCAATTCCTAATCTTTAGTTGAT 720
QY 821 CAGACATGCTACAGGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 880
Db 721 CAGACATGCTACAGGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 780
QY 881 TTTGCGAGCAGAACCGCTAGTGGAGCTGTAGAAAAATGCTAAATCTAATAACAGTATAAGC 940

Handwritten signature/initials.

Db 781 TTTGAGGACGAGACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAAATACAGTATAAGC 840
Qy 941 AACATAGATTACGCTAAAGCAGCAATCGCTAGCTAGACACAAATAGCTGAGCTCAG 1000
Db 841 AACATAGATTACGCTAAAGCAGCAATCGCTAGCTAGACACAAATAGCTGAGCTCAG 900
Qy 1001 AAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAAATACAGGCTCAG 1060
Db 901 AAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAAATACAGGCTCAG 960
Qy 1061 AAGATCTTAAATAATCAAACTCGAGATGGTCTGATGTTCCAAATCCAGGAAGTACA 1120
Db 961 AAGATCTTAAATAATCAAACTCGAGATGGTCTGATGTTCCAAATCCAGGAAGTACA 1020
Qy 1121 GTTGGAGGCTCAAGCAACAGGAAGTAGTATGGTAGTATTCGGTTCCTCATGCTGTTA 1180
Db 1021 GTTGGAGGCTCAAGCAACAGGAAGTAGTATGGTAGTATTCGGTTCCTCATGCTGTTA 1080
Qy 1181 GATGATGCTGAAATGAGACCGCTTCCATTTTGATGCTGGGTTTCGTCAGATGATTCAC 1240
Db 1081 GATGATGCTGAAATGAGACCGCTTCCATTTTGATGCTGGGTTTCGTCAGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATTCCTGATCTCAAGCTGCCCAACAGGAGCTCGACGCAAGCT 1300
Db 1141 ATGTTCAATACGGAATTCCTGATCTCAAGCTGCCCAACAGGAGCTCGACGCAAGCT 1200
Qy 1301 AGAGCAGGAAAGCGCTGGAGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAA 1360
Db 1201 AGAGCAGGAAAGCGCTGGAGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAA 1260
Qy 1361 GCTTTAGAGCGCTCTAGGTAAGCTGGGCAACAAAGGCGATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGCTCTAGGTAAGCTGGGCAACAAAGGCGATCTCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTCTGCTGCTGTGAGCGCAGAGTCTCTCCGCTGAGCGAGTCTTATA 1480
Db 1321 CAGATCGCTCTGCTGCTGTGAGCGCAGAGTCTCTCCGCTGAGCGAGTCTTATA 1380
Qy 1481 GGGTCATCTGTAACAGCTTTTCAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1540
Db 1381 GGGTCATCTGTAACAGCTTTTCAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1440
Qy 1541 CAGATACAGAGGTTATGATGCTTACAAATCCATCATATGATGCTATGTTAGGCGACA 1600
Db 1441 CAGATACAGAGGTTATGATGCTTACAAATCCATCATATGATGCTATGTTAGGCGACA 1500
Qy 1601 AATGATCGACTCTGATGATTAACAAATGATGATACCCCGCTCTCACAGTCCGTT 1660
Db 1501 AATGATCGACTCTGATGATTAACAAATGATGATACCCCGCTCTCACAGTCCGTT 1560
Qy 1661 CCTAGACAGCAACAGAGCTCGAGGACCCAGAAAAACAGATCAAGCCCTCGCTAGGGTG 1720
Db 1561 CCTAGACAGCAACAGAGCTCGAGGACCCAGAAAAACAGATCAAGCCCTCGCTAGGGTG 1620
Qy 1721 ATTCTGCAATAGCAGAACTTTGGAGATGCTATAGTCAAGTTTCGGCACTCAATCT 1780
Db 1621 ATTCTGCAATAGCAGAACTTTGGAGATGCTATAGTCAAGTTTCGGCACTCAATCT 1680
Qy 1781 GTAATGAGATCATCCAGTGAATCTTCAAGCGAATATAGGAGATCAGCAAAAGCTT 1840
Db 1581 GTAATGAGATCATCCAGTGAATCTTCAAGCGAATATAGGAGATCAGCAAAAGCTT 1740
Qy 1841 ACATCGGAGTGACAAAGCCCTCCACAGTTTGGCTATCCTTATGTCAACTTCTTAATGAC 1900
Db 1741 ACATCGGAGTGACAAAGCCCTCCACAGTTTGGCTATCCTTATGTCAACTTCTTAATGAC 1800
Qy 1901 TCTACAGAGATTCTAGCTTAATTAAGAGTTTGTGCTGAAGATCTAGACAGCA 1960
Db 1801 TCTACAGAGATTCTAGCTTAATTAAGAGTTTGTGCTGAAGATCTAGACAGCA 1860
Qy 1961 GCTCAATATAAGCACCTTCTCTTGAACGACCTCTTGTGTTTATTCAGCAGGCTGCTGTC 2020

Db 1861 GCTGAATATAAGCACCTTCTCTTGAAGCAACTCTTGTATTATTACAGAGGTCGTGTC 1920
Qy 2021 AATATCGGCTCTCTATATTTCTGTTATCTTCAATAA 2056
Db 1921 AATATCGGCTCTCTATATTTCTGTTATCTTCAATAA 1956
RESULT 4
ABL92612
ID ABL92612 standard; DNA; 1956 BP.
XX ABL92612;
AC
XX 29-AUG-2003 (revised)
DT 05-JUN-2002 (first entry)
XX
DE Chlamydia pneumoniae DNA sequence SEQ ID NO:385.
XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
KW antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response; gene; ds.
XX
OS Chlamydophila pneumoniae.
XX
XX WD2020208267-A2.
XX 31-JAN-2002.
XX 20-JUL-2001; 2001WO-US023121.
XX 20-JUL-2000; 2000US-00620412.
XX 23-APR-2001; 2001US-00841132.
XX
XX (CORI-1) CORIXA CORP.
XX
XX Fling SP, Skeiky YAM, Probst P, Bhatia A;
XX WPI; 2002-179901/23.
XX
XX Novel compositions comprising Chlamydia Capl protein and its use in the
XX treatment of Chlamydia infection.
XX
XX Disclosure; Page 336-337; 537pp; English.
XX
XX The present invention describes compositions comprising a Chlamydia Capl
XX protein and methods for the diagnosis and therapy of Chlamydia infection.
XX Chlamydia DNA and protein sequences from the present invention can have
XX antibacterial and immunostimulant activities, and can be used in
XX vaccines. Compounds from the present invention can be used for eliciting
XX an immune response, specifically stimulating a Chlamydia-specific T-cell
XX response or inhibiting the development of a Chlamydia infection in an
XX animal. Methods from the present invention can be used: for detecting the
XX presence of Chlamydia in a patient; to stimulate and/or expand T cells
XX specific for a Chlamydia protein; and for treatment of a Chlamydia
XX infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
XX sequences used in the exemplification of the present invention. (Updated
XX on 29-AUG-2003 to standardise OS field)
XX
XX Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;
Query Match 90.7%; Score 1956; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 101 ATGGTTAATCTTATGGTCCAGGTCCTATAGAGAAACAGACGACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTTATGGTCCAGGTCCTATAGAGAAACAGACGACCTCCCGCAGAT 60
Qy 161 CTCTTCTGCTCAAGATTGGAGGCGAGTGCAGCAAAATAGAGTCCGGAAGCTCAAGAATA 220
Db 61 CTCTTCTGCTCAAGATTGGAGGCGAGTGCAGCAAAATAGAGTCCGGAAGCTCAAGAATA 120
Qy 221 GCAGGTGCGGAGCTAAGCCTAAAGAAATCTAGACCGGATTCTGTAGAGCGATGAGCATC 280

Db 121 GCAGTGGGAACTAAGCTAAAGAAATCTAAGACCGATTCGTAGACGATGGAGCATC 180
QY 281 TTGGTCTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGT 340
Db 181 TTGGTCTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGT 240
QY 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGAGCTCAACGACAGCGACCGCACT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGAGCTCAACGACAGCGACCGCACT 300
QY 401 AGGCTCTCTCAACCCAGTGTGATGATTAAGACTCAAGCGCAAAACAGCTTACGATCT 460
Db 301 AGGCTCTCTCAACCCAGTGTGATGATTAAGACTCAAGCGCAAAACAGCTTACGATCT 360
QY 461 ATCTTTTACTCAACATCACTAGTCAATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 520
Db 361 ATCTTTTACTCAACATCACTAGTCAATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 420
QY 521 GTCCTAATATAAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 580
Db 421 GTCCTAATATAAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 480
QY 581 GAACTAGATAGCCGATGCGAGTTAAAGTTGGCGCGAAATTAAGATTAAGGAAATAT 640
Db 481 GAACTAGATAGCCGATGCGAGTTAAAGTTGGCGCGAAATTAAGATTAAGGAAATAT 540
QY 641 GCTTCGGATAACCAAGCGATCTTGACTCTTTAGTAAACTGACTCTCTCGACCTCTTA 700
Db 541 GCTTCGGATAACCAAGCGATCTTGACTCTTTAGTAAACTGACTCTCTCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCTCCAATCTGTAGCAACAAATAACAAAGCAGCTGAGCTCTTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAATCTGTAGCAACAAATAACAAAGCAGCTGAGCTCTTAAAGAG 660
QY 761 ATGCAAGATAACCCAGTGTCCCGGAAAGCGCTGCAATGCTCAATCTTTAGTTCAT 820
Db 661 ATGCAAGATAACCCAGTGTCCCGGAAAGCGCTGCAATGCTCAATCTTTAGTTCAT 720
QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAAGGATCATAT 880
Db 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAAGGATCATAT 780
QY 881 TTTGAGGACGAAACGCTAGTGGAGCTGTAGAAAAATGCTTAATCTTAATACAGTATAAGC 940
Db 781 TTTGAGGACGAAACGCTAGTGGAGCTGTAGAAAAATGCTTAATCTTAATACAGTATAAGC 840
QY 941 AACATAGATTCAGCTAAAGCAGCAATCGCTACTGCTAGACACAAATAGCTGAGCTCAG 1000
Db 841 AACATAGATTCAGCTAAAGCAGCAATCGCTACTGCTAGACACAAATAGCTGAGCTCAG 900
QY 1001 AAAAGTTCCTCCGACTCTCCAATCTTCAAGAGCGGAAACAAATGGTAAATCAGGCTCAG 1060
Db 901 AAAAGTTCCTCCGACTCTCCAATCTTCAAGAGCGGAAACAAATGGTAAATCAGGCTCAG 960
QY 1061 AAAGATCTTAAATAATCAACCTCGAGATGTTCTGATGTTTCCAATCCAGGAACTACA 1120
Db 961 AAAGATCTTAAATAATCAACCTCGAGATGTTCTGATGTTTCCAATCCAGGAACTACA 1020
QY 1121 GTTGGAGGCTCAAGCAACAGGAAGTAGTATTGGTAGTATTTCGTGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGGCTCAAGCAACAGGAAGTAGTATTGGTAGTATTTCGTGTTTCCATGCTGTTA 1080
QY 1181 GATGATGCTGAAAATGAGACCGCTTCCAATTTGATGTTCTGGGTTTTCGTAGATGTTTCAAC 1240
Db 1081 GATGATGCTGAAAATGAGACCGCTTCCAATTTGATGTTCTGGGTTTTCGTAGATGTTTCAAC 1140
QY 1241 ATGTTCAATACGGAATTCCTGATTTCTAAGCTGCCCAACAGGAGCTCGACGACAACT 1300
Db 1141 ATGTTCAATACGGAATTCCTGATTTCTAAGCTGCCCAACAGGAGCTCGACGACAACT 1200
QY 1301 AGACGACGAAAGCCGCTGGAGATGACAGTCTGCTGTCAGGCGCTGCGAGATGCTCAGAAA 1360

Db 1201 AGACGACGAAAGCCGCTGGAGATGACAGTGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1260
QY 1361 GCTTTAGAGCGCTCTAGTAAAGCTGGGCAACAACAGGCGCATACTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGCTCTAGTAAAGCTGGGCAACAACAGGCGCATACTCAATGCTTTAGGA 1320
QY 1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGTTCCTCCGCTGCGACAAAGTTCTATA 1480
Db 1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGTTCCTCCGCTGCGACAAAGTTCTATA 1380
QY 1481 GGGTCATCTGTAAACAGCTTTTCAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1540
Db 1381 GGGTCATCTGTAAACAGCTTTTCAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1440
QY 1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGA 1600
Db 1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGA 1500
QY 1601 AATGATCGACTCTGATGTTGATAAACAATGTAAAGTACCCCGCTCTCACACGATCCGTT 1660
Db 1501 AATGATCGACTCTGATGTTGATAAACAATGTAAAGTACCCCGCTCTCACACGATCCGTT 1560
QY 1661 CCTAGACGACGAAACAGAGCTCGAGGACCGAAGAAACAGATCAAGCCCTCGTAGGGTG 1720
Db 1561 CCTAGACGACGAAACAGAGCTCGAGGACCGAAGAAACAGATCAAGCCCTCGTAGGGTG 1620
QY 1721 ATTTCTGGCAATAGCAGAACTCTTTGGAGATGCTCTATAGTCAAAGTTTGGGCACTACAATCT 1780
Db 1621 ATTTCTGGCAATAGCAGAACTCTTTGGAGATGCTCTATAGTCAAAGTTTGGGCACTACAATCT 1680
QY 1781 GTAAATGAGATCATCCAGTCCGATCCCTCAAGCGAATTAATGAGGAGATCAGACAAAAGCTT 1840
Db 1681 GTAAATGAGATCATCCAGTCCGATCCCTCAAGCGAATTAATGAGGAGATCAGACAAAAGCTT 1740
QY 1841 ACATCGGACGTGACAAAGCCCTCCACAGTCTTGGCTATCTTATGTCGAATCTTAAATCAG 1900
Db 1741 ACATCGGACGTGACAAAGCCCTCCACAGTCTTGGCTATCTTATGTCGAATCTTAAATCAG 1800
QY 1901 TCTACACAGAAATTCATAGCTTAAATAGAAAGTTTGTGCTGAAGGATCTAGGACAGCA 1960
Db 1801 TCTACACAGAAATTCATAGCTTAAATAGAAAGTTTGTGCTGAAGGATCTAGGACAGCA 1860
QY 1961 GCTGAATAAAGACATCTTCTTTGAAACGAACTCTCTGTTTATTCAGCAGGTCGTGTC 2020
Db 1861 GCTGAATAAAGACATCTTCTTTGAAACGAACTCTCTGTTTATTCAGCAGGTCGTGTC 1920
QY 2021 AATATCGGCTCTCTATATCTGTTTATCTCCAATAA 2056
Db 1921 AATATCGGCTCTCTATATCTGTTTATCTCCAATAA 1956

RESULT 5

ABL91190
ID ABL91190 standard; DNA; 1956 BP.
XX ABL91190;
AC ABL91190;
XX 29-AUG-2003 (revised)
DT 29-JUL-2002 (first entry)
XX
XX Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.
XX
KW Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.
KW human respiratory infection; antigen; immunogen; vaccine; diagnosis;
KW human respiratory disease; cardiovascular disease; atherosclerosis;
KW coronary artery disease; carotid artery stenosis; myocardial infarction;
KW cerebrovascular disease; aortic aneurysm; claudication; stroke;
KW strain CWL029; open reading frame; ORF; gene; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 1..1956
FT /*tag= a

FT XX
PN XX
XX XX
PD XX
PF XX
PR XX
PR 11-JUL-2000; 2000GB-00016363.
PR 21-JUL-2000; 2000GB-00017047.
PR 07-AUG-2000; 2000GB-00017983.
PR 18-AUG-2000; 2000GB-00019368.
PR 14-SEP-2000; 2000GB-00020440.
PR 10-NOV-2000; 2000GB-00022583.
PR 22-DEC-2000; 2000GB-00031706.
XX (CHIR-) CHIRON SPA.
XX Ratti G, Grandi G;
XX WPI; 2002-154726/20.
DR N-PSDB; ABB90532.
XX
PT Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
PT preferably Chlamydia pneumoniae, and for diagnostic purposes.
PS Claim 5; Page 47-48; 364pp; English.
XX
CC Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
CC pneumoniae (strain CWL029), and ABL9184-ABL91373 represent DNA encoding
CC them. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as
CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in the detection of
CC Chlamydia pneumoniae gene expression. The present sequence represents a
CC specifically claimed DNA which encodes a Chlamydia pneumoniae protein of
CC the invention. (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 1956 BP; 506 A; 436 C; 441 G; 473 T; 0 U; 0 Other;

Query Match 90.7%; Score 1956; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGTTTAACTCTATGTCCTCAGTCCCTATAGACGAAACAGACACACCTCCCGCAGAT 160
Db 1 ATGTTTAACTCTATGTCCTCAGTCCCTATAGACGAAACAGACACACCTCCCGCAGAT 60

QY 161 CTTTCTCTCAAGGATTGGAGGCGAGTGCGAGCAATTAAGATGGGAGCTCAAGAATA 220
Db 61 CTTTCTCTCAAGGATTGGAGGCGAGTGCGAGCAATTAAGATGGGAGCTCAAGAATA 120

QY 221 GCAGGTGGGAAGCTAAGCCCTAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATC 280
Db 121 GCAGGTGGGAAGCTAAGCCCTAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATC 180

QY 281 TTGCGTTCTCGAGTGAATGCTCATGAGTCTGCGAGATAGCTGGGTATGCTTCTAGT 340
Db 181 TTGCGTTCTCGAGTGAATGCTCATGAGTCTGCGAGATAGCTGGGTATGCTTCTAGT 240

QY 341 AACAGCTCGTCTTCTACTAGCAGATCTGAGAGCTGGACCTCAACGACAGCGCCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGAGAGCTGGACCTCAACGACAGCGCCGACCT 300

QY 401 ACGCTCTCTCCACCCACGTTTGATGATTATAGACTCAAGCGCAACAGCTTACGATACT 460
Db 301 ACGCTCTCTCCACCCACGTTTGATGATTATAGACTCAAGCGCAACAGCTTACGATACT 360

QY 461 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCT 520
Db 361 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCT 420

QY 521 GTCACTAATATAAGGATACAGCGGCTACTGATAGGAAACCGCAATCGCTGCGGAGTGG 580
Db 421 GTCACTAATATAAGGATACAGCGGCTACTGATAGGAAACCGCAATCGCTGCGGAGTGG 480

QY 581 GAATACTAAGATGCGATGAGTGAAGTTGAGGCGCAATTAAGAAATAGCAATATAT 640
Db 481 GAATACTAAGATGCGATGAGTGAAGTTGAGGCGCAATTAAGAAATAGCAATATAT 540

QY 641 GCTTCGGATAACCAAGCGATTCTTGACTCTTTAGSTAACTGACTCTCTTCGACTCTTA 700
Db 541 GCTTCGGATAACCAAGCGATTCTTGACTCTTTAGSTAACTGACTCTCTTCGACTCTTA 600

QY 701 CAGCTCTCTCTCTCCAACTCTGTAGCAAAACAATAACAGAGCTGAGCTTCTTAAGAG 760
Db 601 CAGCTCTCTCTCTCCAACTCTGTAGCAAAACAATAACAGAGCTGAGCTTCTTAAGAG 660

QY 761 ATGCAAGATAACCCAGTAGTCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTAGTCTCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGAT 720

QY 821 CAGACAGATCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGATGCATAT 880
Db 721 CAGACAGATCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGATGCATAT 780

QY 881 TTTTGCAGGACAGACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAGC 940
Db 781 TTTTGCAGGACAGACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAGC 840

QY 941 AACATAGATTCAAGTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTCAAGTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900

QY 1001 AAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 1060
Db 901 AAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 960

QY 1061 AAAGATCTTAAATAATATCAACCTGAGAGTGTCTGATGCTTCCAAATCCAGGACTACA 1120
Db 961 AAAGATCTTAAATAATATCAACCTGAGAGTGTCTGATGCTTCCAAATCCAGGACTACA 1020

QY 1121 GTTGGAGGCTCCAAAGCAACAAAGGAAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTA 1180
Db 1021 GTTGGAGGCTCCAAAGCAACAAAGGAAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTA 1080

QY 1181 GATGATGCTGAAATGAGACCGCTTCCAAATTTGATGCTGGGTTTCTGATGATGATTCAC 1240
Db 1081 GATGATGCTGAAATGAGACCGCTTCCAAATTTGATGCTGGGTTTCTGATGATGATTCAC 1140

QY 1241 ATGTTCAATACGAAATCTCTGATTTCTCAAGCTGCCCAACAGGAGCTCCGAGCACAAGCT 1300
Db 1141 ATGTTCAATACGAAATCTCTGATTTCTCAAGCTGCCCAACAGGAGCTCCGAGCACAAGCT 1200

QY 1301 AGACGACGAAAGCCCGCTGGAGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAA 1360
Db 1201 AGACGACGAAAGCCCGCTGGAGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAA 1260

QY 1361 GCTTTAGAGCGGCTCTAGGTAAGCTGGGCAACAAAGGCAATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGGCTCTAGGTAAGCTGGGCAACAAAGGCAATCTCAATGCTTTAGGA 1320

QY 1421 CAGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1480
Db 1321 CAGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380


```
Db      541  |||||GCTTCGGATAACCAAGCGATCTTTGACTCTTTAGGTAACTGACTTCTCTCGACCTCTTA 600
Qy      701  CAGGCTGCTCTCTCCATCTGTGAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG 760
Db      601  |||||CAGGCTGCTCTCTCCATCTGTGAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG 660
Qy      761  ATGCAAGATAACCCAGTAGTCCCAGGGGAAACGGCTGCAATTTGCTCAATCTTTAGTTGAT 820
Db      661  ATGCAAGATAACCCAGTAGTCCCAGGGGAAACGGCTGCAATTTGCTCAATCTTTAGTTGAT 720
Qy      821  CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATAT 880
Db      721  CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATAT 780
Qy      881  TTTTCAGACAGAACGCTAGTGGAGCTGTAGAAATGCTAAATCTAAATCAAGTATAGC 940
Db      781  TTTTCAGACAGAACGCTAGTGGAGCTGTAGAAATGCTAAATCTAAATCAAGTATAGC 840
Qy      941  AACATAGATTGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 1000
Db      841  AACATAGATTGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 900
Qy      1001  AAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 1060
Db      901  AAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 960
Qy      1061  AAAGATCTTAAATAATCAAACTGCGAGATGTTCTGATGTTCCAAATCCAGGAACATACA 1120
Db      961  AAAGATCTTAAATAATCAAACTGCGAGATGTTCTGATGTTCCAAATCCAGGAACATACA 1020
Qy      1121  GTTGGAGGCTTCAAGCAACAAGGAAGTAGTATTGGTAGTAGTTCGTGTTTCCATGCTGTTA 1180
Db      1021  GTTGGAGGCTTCAAGCAACAAGGAAGTAGTATTGGTAGTAGTTCGTGTTTCCATGCTGTTA 1080
Qy      1181  GATGATGCTGAAATGAGACCGCTTCCATTTTGATGTCGGGTTTGGTACAGATGATTCAC 1240
Db      1081  GATGATGCTGAAATGAGACCGCTTCCATTTTGATGTCGGGTTTGGTACAGATGATTCAC 1140
Qy      1241  ATGTTCAATACGGAATTCCTGATTCTCAAGCTGCCCAACAGGAGCTCGCAGACCAAGCT 1300
Db      1141  ATGTTCAATACGGAATTCCTGATTCTCAAGCTGCCCAACAGGAGCTCGCAGACCAAGCT 1200
Qy      1301  AGACGAGCAAAAGCCGCTGGAGATGACAGATGCTGCTGAGCGCTGGCAGATGCTCAGAAA 1360
Db      1201  AGACGAGCAAAAGCCGCTGGAGATGACAGATGCTGCTGAGCGCTGGCAGATGCTCAGAAA 1260
Qy      1361  GCTTTAGAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGSCATCACTCAATGCTTTAGGA 1420
Db      1261  GCTTTAGAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGSCATCACTCAATGCTTTAGGA 1320
Qy      1421  CAGATCGCTTCTGCTGCTGTGTGAGCGCGCAGAGTTCTCTCCGCTGCGAGCAAGTTCTATA 1480
Db      1321  CAGATCGCTTCTGCTGCTGTGTGAGCGCGCAGAGTTCTCTCCGCTGCGAGCAAGTTCTATA 1380
Qy      1481  GGGTCATCTGTAAACAGCTTTTACAAGCTTCAAGATCTACAGGTTCTGATTATATAAACA 1540
Db      1381  GGGTCATCTGTAAACAGCTTTTACAAGCTTCAAGATCTACAGGTTCTGATTATATAAACA 1440
Qy      1541  CAGATATCAGCAGGTTATGATGCTTACAAATCCATGATGCTATGATGGGACGGA 1600
Db      1441  CAGATATCAGCAGGTTATGATGCTTACAAATCCATGATGCTATGATGGGACGGA 1500
Qy      1601  AATGATGCGACTCGTGATGTATAAACAAATGATGATGATGATGATGATGATGATGATGAT 1660
Db      1501  AATGATGCGACTCGTGATGTATAAACAAATGATGATGATGATGATGATGATGATGATGAT 1560
Qy      1661  CCTAGAGCAAGAACAGAGCTCGAGGACAGAAAAAACAAGATCAAGCCCTCGCTAGGGTG 1720
Db      1561  CCTAGAGCAAGAACAGAGCTCGAGGACAGAAAAAACAAGATCAAGCCCTCGCTAGGGTG 1620
Qy      1721  ATTTCTGCAATAGCAGAACTCTGGAGATGCTTATAGTCAAGTTTGGGCACTACAATCT 1780
```

```
Db      1621  ATTTCTGGCAATAGCAGAACTCTTTGGAGATGTTCTATAGTCAAGTTTCGGCACTACAATCT 1680
Qy      1781  GTATTCAGATCATCCAGTCGAATCTCTCAAGCGAATAATAGGAGATCAGACAAAAGCTT 1840
Db      1681  GTATTCAGATCATCCAGTCGAATCTCTCAAGCGAATAATAGGAGATCAGACAAAAGCTT 1740
Qy      1841  ACATCGSCAGTGACAAAAGCCTCCACAGTTTGGGCTATCTCTATGTGCAACTTTCTAATGAC 1900
Db      1741  ACATCGSCAGTGACAAAAGCCTCCACAGTTTGGGCTATCTCTATGTGCAACTTTCTAATGAC 1800
Qy      1901  TCTACACAGAAAGTTCTATAGCTAAATAGAAAGTTTGTGCTGCTGAAGATCTAGGACAGCA 1960
Db      1801  TCTACACAGAAAGTTCTATAGCTAAATAGAAAGTTTGTGCTGCTGAAGATCTAGGACAGCA 1860
Qy      1961  GCTGMAATAAAGCACTTCTCTTGTGAACGAACCTCTCTGTTTATTCAGCAGGTCCTGGTC 2020
Db      1861  GCTGMAATAAAGCACTTCTCTTGTGAACGAACCTCTCTGTTTATTCAGCAGGTCCTGGTC 1920
Qy      2021  AATATCGGCTCTCTATATATCTGTTATCTCTCAATAA 2056
Db      1921  AATATCGGCTCTCTATATATCTGTTATCTCTCAATAA 1956
```

RESULT 7

```
AA0202064
ID      AA0202064 standard; DNA; 1852 BP.
XX
AC      AAD02064;
XX
DT      26-MAR-2001 (first entry)
XX
DE      5'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.
XX
KW      76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
XX      upper respiratory tract disease; bronchitis; sinusitis;
XX      acute respiratory disease; cough; sore throat; hoarseness; fever;
XX      vaccine; immunisation; treatment; truncation mutant; ds.
XX
OS      Chlamydia pneumoniae.
XX
FH      Key Location/Qualifiers
FT      CDS 1..1752
FT      /*tag= a
FT      /product= "5'-truncated Chlamydia pneumoniae 76kDa
FT      protein"
FT      /transl_except= (pos:1489..1491, aa:Ile)
```

MO200066739-A2.

09-NOV-2000.

03-MAY-2000; 2000WO-CA000511.

03-MAY-1999; 99US-0132270P.

30-JUN-1999; 99US-0141276P.

(AVET) AVENTIS PASTEUR LTD.

Murdin AD, Comen RP, Wang J, Dunn P;

WPI; 2000-687542/67.

P-PSDB; AAY71955.

Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful for vaccinating against Chlamydia infections.

Claim 2b; Page 97-99; 112pp; English.

The present sequence is a DNA coding for 5'-truncated Chlamydia pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory tract disease, bronchitis, sinusitis and

CC acute respiratory disease such as cough, sore throat, hoarseness, fever;
CC and abnormal chest sounds on auscultation). C. pneumoniae sequence is
CC also used as vaccines for immunising humans against diseases caused by C.
CC pneumoniae
XX
SQ

Query Match 85.8%; Score 1848.8; DB 3; Length 1852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 305 ATGACTCTGGCAGATAAGCTGGTATTCCTCTAGTAACACGCTGCTCTCTACTAGCAGA 364
Db 1 ATGAGTCTGGCAGATAAGCTGGTATTCCTCTAGTAACACGCTGCTCTCTACTAGCAGA 60
QY 365 TCTGCAGACGCTGGATCTCAACACACGACGACCGCACCTACGCTCTCCACCCACGCTTTGAT 424
Db 61 TCTGCAGACGCTGGATCTCAACACACGACGACCGCACCTACGCTCTCCACCCACGCTTTGAT 120
QY 425 GATTATAGACTCAAGCGGAAACAGCTTACGATACATCTATCTTACCTCAACATCACTAGCT 484
Db 121 GATTATAGACTCAAGCGGAAACAGCTTACGATACATCTATCTTACCTCAACATCACTAGCT 180
QY 485 GACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGTCACTATATATAAGATACAGCG 544
Db 181 GACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGTCACTATATATAAGATACAGCG 240
QY 545 GCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGCACTT 604
Db 241 GCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGCACTT 300
QY 605 AAAGTTGGCGGCAAAATPACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATCTT 664
Db 301 AAAGTTGGCGGCAAAATPACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATCTT 360
QY 665 GACTCTTTAGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCATCTGTA 724
Db 361 GACTCTTTAGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCATCTGTA 420
QY 725 GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTAGTCCCA 784
Db 421 GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTAGTCCCA 480
QY 785 GGGAAACGCTGCAATTTGCTCAATCTTTAGTGAATGATGATGATGATGATGATGATGAT 844
Db 481 GGGAAACGCTGCAATTTGCTCAATCTTTAGTGAATGATGATGATGATGATGATGATGAT 540
QY 845 ATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTTCAGGACAGACGCTAGTGGG 904
Db 541 ATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTTCAGGACAGACGCTAGTGGG 600
QY 905 GCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTTAAAGCAGCA 964
Db 601 GCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTTAAAGCAGCA 660
QY 965 ATCGCTACTGTAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGAGCTCTCCAAAT 1024
Db 661 ATCCCTACTGTAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGAGCTCTCCAAAT 720
QY 1025 CTTCAAGAAGCGGAACAAATGGTAAATACAGGCTGAGAAAAGATCTTAAAAATATCAAACT 1084
Db 721 CTTCAAGAAGCGGAACAAATGGTAAATACAGGCTGAGAAAAGATCTTAAAAATATCAAACT 780
QY 1085 GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACAGGA 1144
Db 781 GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACAGGA 840
QY 1145 AGTAGTATTTGCTAGTATTCGTTTCCATGCTGTAGATGATGATGATGATGATGATGATG 1204
Db 841 AGTAGTATTTGCTAGTATTCGTTTCCATGCTGTAGATGATGATGATGATGATGATGATG 900
QY 1205 TCCATTTTGTGTTTGGTCTGCTGATGATTCACATGTTCAATACGGAATTCCTGAT 1264

Db 901 TCCATTTTGTGTTCTGGTTCGTCAGATGATTCACATGTTCAATACGGAATTCCTGAT 960
QY 1265 TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAGCTAGACAGGAAAGCGCTGGAGAT 1324
Db 961 TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAGCTAGACAGGAAAGCGCTGGAGAT 1020
QY 1325 GACAGTCTGCTGACGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTTAGGTTAA 1384
Db 1021 GACAGTCTGCTGACGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTTAGGTTAA 1080
QY 1385 GCTGGGCAACACAGGCGATCTCAATGCTTTTAGGAAGATGCTGCTGCTGCTGCTGCTG 1444
Db 1081 GCTGGGCAACACAGGCGATCTCAATGCTTTTAGGAAGATGCTGCTGCTGCTGCTGCTG 1140
QY 1445 AGCGCAGAGTTCTCCCGCTGCGCAGCAAGTTCTATAGSGTCATCTGTAAACACAGCTTTAC 1504
Db 1141 AGCGCAGAGTTCTCCCGCTGCGCAGCAAGTTCTATAGSGTCATCTGTAAACACAGCTTTAC 1200
QY 1505 AAGACCTCAAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATGATGCT 1564
Db 1201 AAGACCTCAAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
QY 1565 TACAAATCCATCAATGATGCTATGCTAGGCGACGAAATGATGCGACTCTGTGATGATGA 1624
Db 1261 TACAAATCCATCAATGATGCTATGCTAGGCGACGAAATGATGCGACTCTGTGATGATGA 1320
QY 1625 AACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1684
Db 1321 AACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1685 GGACGAGAAAACACAGATCAAGCCCTCGCTAGGTTGATTTCTGSCAATAGCAGAACTCTT 1744
Db 1381 GGACGAGAAAACACAGATCAAGCCCTCGCTAGGTTGATTTCTGSCAATAGCAGAACTCTT 1440
QY 1745 GGAGATGCTATAGTCAAGTTTCGCGACTACATCTGTANTGAGATCACTCCAGTCCAAAT 1804
Db 1441 GGAGATGCTATAGTCAAGTTTCGCGACTACATCTGTANTGAGATCACTCCAGTCCAAAT 1500
QY 1805 CCTCAACGGAATATGAGGAGATCAGACAAAAGCTTACATCGGCGAGTGACAAAAGCCCTCCA 1864
Db 1501 CCTCAACGGAATATGAGGAGATCAGACAAAAGCTTACATCGGCGAGTGACAAAAGCCCTCCA 1560
QY 1865 CAGTTTGGCTATGCTTAAAGTCAACTTCTPAATGACTCTACACAGAAAGTTCAATAGCTAAA 1924
Db 1561 CAGTTTGGCTATGCTTAAAGTCAACTTCTPAATGACTCTACACAGAAAGTTCAATAGCTAAA 1620
QY 1925 TTAGAAAAGTTTGTGCTGAAGGATCTAGCAGCAGCTGAAATATAAAGCAGCTTTCCCTTT 1984
Db 1521 TTAGAAAAGTTTGTGCTGAAGGATCTAGCAGCAGCTGAAATATAAAGCAGCTTTCCCTTT 1680
QY 1985 GAAAACGAACTCTTGTGTTTATTCAGCAGGTCGCTGCTCAATATCGGCTCTCTATATTCGCT 2044
Db 1681 GAAAACGAACTCTTGTGTTTATTCAGCAGGTCGCTGCTCAATATCGGCTCTCTATATTCGCT 1740
QY 2045 TATCTCCAAATAACACACCTCAAGTTCGTTTGGAGAGATATATGCTGCTTGGTAAAG 2104
Db 1741 TATCTCCAAATAACACACCTCAAGTTCGTTTGGAGAGATATATGCTGCTTGGTAAAG 1800
QY 2105 CTTTGTGTTGAGGCTTACCAACACACCTAGAACGATCTTCAATAATAAAGA 2156
Db 1801 CTTTGTGTTGAGGCTTACCAACACACCTAGAACGATCTTCAATAATAAAGA 1852

RESULT 8

AAD02066

ID AAD02066 standard; DNA; 2238 BP.

XX AAD02066;

AC AAD02066;

XX 26-MAR-2001 (first entry)

XX C. pneumoniae 76 kDa protein truncation mutant fusion gene.

XX

76 kDa protein; bactericidal; diagnosis; prevention; pneumonia; upper respiratory tract disease; bronchitis; sinusitis; acute respiratory disease; cough; sore throat; hoarseness; fever; vaccine; immunisation; treatment; truncation mutant; fusion gene; ds. Chlamydophila pneumoniae. Synthetic.

Key Location/Qualifiers
misc_feature 1..665
/tag= a
/note= "This part of the sequence is unrelated to C. pneumoniae 76 kDa gene"
CDS 766..2238
/tag= b
/product= "Truncated Chlamydia pneumoniae 76kDa protein"
misc_feature 2122..2238
/tag= c
/note= "This part of the sequence is unrelated to C. pneumoniae 76 kDa gene"

WO20066739-A2.
09-NOV-2000.
03-MAY-2000; 2000WO-CA000511.
03-MAY-1999; 99US-0132270P.
30-JUN-1999; 99US-0141276P.
(AVET) AVENTIS PASTEUR LTD.
Murdin AD, Oomen RP, Wang J, Dunn P;
WPI; 2000-687542/67.
P-PSDB; RAY71957.
Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful for vaccinating against Chlamydia infections.
Claim 32; Fig 3; 112pp; English.
The present sequence is a DNA coding for a fusion protein comprising a truncated Chlamydia pneumoniae 76 kDa protein and vector-encoded residues. C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory tract disease, bronchitis, sinusitis and acute respiratory disease such as cough, sore throat, hoarseness, fever; and abnormal chest sounds on auscultation). C. pneumoniae sequence is also used as vaccines for immunising humans against diseases caused by C. pneumoniae

Query Match 68.0%; Score 1465.8; DB 3; Length 2238;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1478; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Sequence 2238 BP; 689 A; 459 C; 497 G; 593 T; 0 U; 0 Other;

1 ATAAATCTTAAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 60
666 ATAAATCTTAAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 725
61 ATAAATCTTAAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 120
726 ATAAATCTTAAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 785
121 AGGTCCTATAGCAAGACAGACGACCTCCGAGAGCTTTCTCTCAAGATTGGA 180
786 AGGTCCTATAGCAAGACAGACGACCTCCGAGAGCTTTCTCTCAAGATTGGA 845
181 GCGAGTGCGAGCAATTAAGAGTGGGAGGCTCAAGAGTACGAGTGGGAGGCTTAAGCC 240
846 GCGAGTGCGAGCAATTAAGAGTGGGAGGCTCAAGAGTACGAGTGGGAGGCTTAAGCC 905

241 TAAAGATCTAAGACCGATTTCTGTAGAGCGATGAGCATCTTGGTTCGAGTGAATGC 300
906 TAAAGATCTAAGACCGATTTCTGTAGAGCGATGAGCATCTTGGTTCGAGTGAATGC 965
301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTACTAG 360
966 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTACTAG 1025
361 CAGATCTGAGAGCGTGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCGAGTT 420
1026 CAGATCTGAGAGCGTGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCGAGTT 1085
421 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATACCTATCTTTACCTCAACATCACT 480
1086 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATACCTATCTTTACCTCAACATCACT 1145
481 AGCTGACATACAGGCTCTTTGGTGGAGCTCCAGGATGCTGTCACTAATAATAAGGATAC 540
1146 AGCTGACATACAGGCTCTTTGGTGGAGCTCCAGGATGCTGTCACTAATAATAAGGATAC 1205
541 AGCGCTACTGATGAGCAACCGCAATCGCTGGAGTGGGAACTAAGATGCCGATGC 600
1206 AGCGCTACTGATGAGCAACCGCAATCGCTGGAGTGGGAACTAAGATGCCGATGC 1265
601 AGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGAT 660
1266 AGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGAT 1325
661 TCTTGACTCTTTAGGTAACTGACTTCTTCGAGCTCTTTACAGCTCTCTTCTCCAATC 720
1326 TCTTGACTCTTTAGGTAACTGACTTCTTCGAGCTCTTTACAGCTCTCTTCTCCAATC 1385
721 TGTAGCAACAATAAACAAGCAGCTGAGCTTTTAAAGAGATCAAGATTAACCAAGCGAT 780
1386 TGTAGCAACAATAAACAAGCAGCTGAGCTTTTAAAGAGATCAAGATTAACCAAGCGAT 1445
781 CCCAGGAAACCGCTGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
1446 CCCAGGAAACCGCTGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 1505
841 ACAGATAGAAAGATGGAAATCGATTAGGATGCAATATTTTCAGGACAGAAACGCTAG 900
1506 ACAGATAGAAAGATGGAAATCGATTAGGATGCAATATTTTCAGGACAGAAACGCTAG 1565
901 TGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATGATTTCAGTAAAGC 960
1566 TGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATGATTTCAGTAAAGC 1625
961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
1626 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1685
1021 AATTCTTCAAGAGCGGACAAATGTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
1686 AATTCTTCAAGAGCGGACAAATGTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1745
1081 ACCTGAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1140
1746 ACCTGAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1805
1141 AGGAAGTAGTATTGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGCAAAATGAGAC 1200
1806 AGGAAGTAGTATTGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGCAAAATGAGAC 1865
1201 CGCTTCCATTTGATGCTCGGGTTTCGTAGATGATTCATGTTCAATACGGAATATCC 1260
1866 CGCTTCCATTTGATGCTCGGGTTTCGTAGATGATTCATGTTCAATACGGAATATCC 1325
1261 TGATTTCTCAAGCTGCCCAACAGGAGCTCCGAGCACAAGCTAGAGCGGAAAGCCCTGG 1320
1926 TGATTTCTCAAGCTGCCCAACAGGAGCTCCGAGCACAAGCTAGAGCGGAAAGCCCTGG 1985

QY	1321	AGATGACAGTCTGCTGCACGCGTGGCAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGG	1381
DB	1986	AGATGACAGTCTGCTGCACGCGTGGCAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGG	2045
QY	1381	TAAAGCTGGGCAACAACAGAGGTCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT	1440
DB	2046	TAAAGCTGGGCAACAACAGAGGTCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT	2105
QY	1441	TGTGAGCGCAGGAGTTCTCTCCGCTGCAGCAAGTTCTATAG	1481
DB	2106	TGTGAGCGCAGGAG- TACTCCGCTGCAGCAAGTTCTATGG	2145
RESULT 9			
AAD02065			
ID	AAD02065 standard; DNA; 1456 BP.		
XX	AAD02065;		
XX	26-MAR-2001 (first entry)		
XX	3'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.		
XX	76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;		
KW	upper respiratory tract disease; bronchitis; sinusitis;		
KW	acute respiratory disease; cough; sore throat; hoarseness; fever;		
KW	vaccine; immunisation; treatment; truncation mutant; ds.		
XX	Chlamydia pneumoniae.		
OS	Synthetic.		
XX	Key		
XX	CDS		
XX	Location/Qualifiers		
FT	101..1456		
FT	/*tag= a		
FT	/product= "3'-truncated Chlamydia pneumoniae 76KDa		
FT	protein"		
FT	/note= "The coding region does not include stop codon"		
FT	/partial		
XX	WO200066739-A2.		
XX	09-NOV-2000.		
XX	03-MAY-2000; 2000WO-CA000511.		
XX	03-MAY-1999; 99US-0132270P.		
XX	30-JUN-1999; 99US-0141276P.		
XX	(AVET) AVENTIS PASTEUR LTD.		
XX	Murdin AD, Oomen RP, Wang J, Dunn P;		
XX	WPI; 2000-687542/67.		
XX	P-PSDB; AAY71956.		
XX	Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful		
XX	for vaccinating against Chlamydia infections.		
XX	Claim 2c; Page 102-104; 112pp; English.		
XX	The present sequence is a DNA coding for 3'-truncated Chlamydia		
XX	pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the		
XX	diagnosis, prevention and treatment of C. pneumoniae infections (e.g.		
XX	pneumonia, upper respiratory tract disease, bronchitis, sinusitis and		
XX	acute respiratory disease such as cough, sore throat, hoarseness, fever;		
XX	and abnormal chest sounds on auscultation). C. pneumoniae sequence is		
XX	also used as vaccines for immunising humans against diseases caused by C.		
XX	pneumoniae		
XX	Sequence 1456 BP; 452 A; 308 C; 331 G; 365 T; 0 U; 0 Other;		
XX	Query Match		
XX	Best Local Similarity 100.0%; Pred. No. 0;		
XX	97.5%; Score 1455; DB 3; Length 1456;		

Matches	1455;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
1	QY	1	ATAAAATCTTTAAAAACAGGCTCGCAATTAATTAATTAGTCAGAGAGCTTTTTTTTTTATTTTTT	60					
1	Db	1	ATAAAATCTTTAAAAACAGGCTCGCAATTAATTAATTAGTCAGAGAGCTTTTTTTTTTATTTTTT	60					
61	QY	61	ATAATAAAACATAAAAGATTTTTTATTTATTTTTTGGAGTTTTTATGGTTAAATCCTTATGGTCC	120					
61	Db	61	ATAATAAAACATAAAAGATTTTTTATTTATTTTTTGAGTTTTTATGGTTAAATCCTTATGGTCC	120					
121	QY	121	AGGTCCTATAGACGAAACAGAACCGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA	180					
121	Db	121	AGGTCCTATAGACGAAACAGAACCGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA	180					
181	QY	181	GCGGAGTCGACGAATAAGAGTGGCGGAGCTCAAGAANTAGCAGAGTGGCGAAGCTTAAGCC	240					
181	Db	181	GCGGAGTCGACGAATAAGAGTGGCGGAGCTCAAGAANTAGCAGAGTGGCGAAGCTTAAGCC	240					
241	QY	241	TAAAGAAATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTGGTCTTGCAGTGAATGC	300					
241	Db	241	TAAAGAAATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTGGTCTTGCAGTGAATGC	300					
301	QY	301	TCTCATGAGTCTGGCAGATAAGCTGGGTAATGGTTCTAGTAACAGCTCGCTTCTTACTAG	360					
301	Db	301	TCTCATGAGTCTGGCAGATAAGCTGGGTAATGGTTCTAGTAACAGCTCGCTTCTTACTAG	360					
361	QY	361	CAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACCTACGCTCTCTCCACCCACGTT	420					
361	Db	361	CAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACCTACGCTCTCTCCACCCACGTT	420					
421	QY	421	TGATGATTATAAGACTCAAGCGGAAACAGCTTAGATATACTTTTACCTCAACATCACT	480					
421	Db	421	TGATGATTATAAGACTCAAGCGGAAACAGCTTAGATATACTTTTACCTCAACATCACT	480					
481	QY	481	AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACATAATAAAGGATAC	540					
481	Db	481	AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACATAATAAAGGATAC	540					
541	QY	541	AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATCCGATGC	600					
541	Db	541	AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATCCGATGC	600					
601	QY	601	AGTTAAAGTTGGCGGCAAAATACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGAT	660					
601	Db	601	AGTTAAAGTTGGCGGCAAAATACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGAT	660					
661	QY	661	TCTTGACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC	720					
661	Db	661	TCTTGACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC	720					
721	QY	721	TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGGAAGATAACCCAGTAGT	780					
721	Db	721	TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGGAAGATAACCCAGTAGT	780					
781	QY	781	CCGAGGGAAAACCCCTGCAATTCGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC	840					
781	Db	781	CCGAGGGAAAACCCCTGCAATTCGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC	840					
841	QY	841	ACAGATAGAAAAGATGGAAAATCGGATTAGGGATGCAATTTTTGCAGGACAGAAACGCTAG	900					
841	Db	841	ACAGATAGAAAAGATGGAAAATCGGATTAGGGATGCAATTTTTGCAGGACAGAAACGCTAG	900					
901	QY	901	TGGAGCTGTAGAAAATGCTTAATCTAATACAGTATAAGCAACATAGATTTCAGCTAAAGC	960					
901	Db	901	TGGAGCTGTAGAAAATGCTTAATCTAATACAGTATAAGCAACATAGATTTCAGCTAAAGC	960					
961	QY	961	AGCAATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC	1020					
961	Db	961	AGCAATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC	1020					
1021	QY	1021	AATCTTCAAGACGCGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAA	1080					
1021	Db	1021	AATCTTCAAGACGCGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAA	1080					

Db	5498	TTGAGGCTCAGCTGA---CTGCGATGTCAGATCAA	CTGTTGTTGGGATGGCGAGCTCC	5554
Qy	1347	CAGATGCTCAGAAAGCTTTAGAACGGCTCTAGTAAAGCTGGCGCAACAAC	-----	1397
Db	5555	CAGCGGAAATACAGCAANTCAAGATGCTCTTGGCAAGCTTTTGAAACAACCATCAGCAG	5614	
Qy	1398	AGGCGATACTCAATGCTTTTAGGCAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGTTC	1457	
Db	5615	ATGGTTAGCTACAGCTATGGACAAGTGGCTTTTGAGCTGCCAAGTTTGGAGGAGGCT	5674	
Qy	1458	CTCCCGTGCAGCAAGTTCTATAGGGTCACTCTGTAAACAGCTTTTACAAGACCTCAAAAT	1517	
Db	5675	CCGCAGGAACAGCTGGCACTGTCCAGATGAATGTAAACAGCTTTTACAAGACAGCGTTTT	5734	
Qy	1518	CTACAGTTCT-----GATTATAAACACAGATATCAGCAGGTTATGATGCTTACAAAT	1571	
Db	5735	CTTCGACTCTTCAAGCTCTTATCAGCAGCAGCTTCCGATGGATATCTGCTTTACAAA	5794	
Qy	1572	CCATCAATGATGCCCTATGTAGGCGACGAATGATGCCACTCGCTGATGTGATAAACAAATG	1631	
Db	5795	CACTGAACCTTTATATTCGAAAGCAGAGCGCGTGCAGTCAGCTATTAGTCAAACTG	5854	
Qy	1632	TAAGTACCCCGCTCTCACAGATCCGTTCTAGAGCACGAACAGAGCTCGAGGACCG	1691	
Db	5855	CAAAATCCCGCTTTCCAGAAAGCGTTTCTCGTTCTGGCATAGAAAGTCAAGSACGCACTG	5914	
Qy	1692	AAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGAACTCTTTGGAGATG	1751	
Db	5915	CAGATGCTAGCCAAAGACGACGAAACTATTGTGAGAGATAGCCAAACGTTAGGTGATG	5974	
Qy	1752	TCTATAGTCAAGTTTGGSCACTCAACTCTGTAATGCGAGTCAATCCAGTCCGAATCCTCAAG	1811	
Db	5975	TATATAGCCGCTTACAGTTCAGATTCTTTGATGCTACGATTGTGAGCAATCCGCAAG	6034	
Qy	1812	CGAATAATAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCCTCCACAGTTTG	1871	
Db	6035	CAAAATCAAGAAGAGATTATGAGAAGCTCACGGCATCTATTAGCAAAAGCTCCACAATTTG	6094	
Qy	1872	GCTATCCTTATGTGCAACTTTCTAATGACTCTACAGAAAGTTCATAGCTAAATAGAAA	1931	
Db	6095	GGTATCCTGCTGTTTCAAGAAATCTGTGGATAGCTTGCAGAAATTTGCTGCACAAATTTGAAA	6154	
Qy	1932	GTTTGTTTGTCTGAAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTCTTTTGAACGA	1991	
Db	6155	GAGAGTTTGTGTATGGGAAAGTAGTCTCGCAGAAATCTCAAGAGANTGCGTTTGAAGAAC	6214	
Qy	1992	ACTCCTTTTATTTACAGCAGGTGTGGTCAATATCGGCTCTCTATATTCTGGTTATCTCC	2051	
Db	6215	AGCCGCTTTCATTCACAGGTGTGGTAACAATGTCTCTCTATCTCTGCTTATCTTT	6274	
Qy	2052	AATAAC	2057	
Db	6275	CTTAAC	6280	

RESULT 12

RESUL 12
AAA63292

AAA63292
ID AAA63292 standard: DNA: 1537 BP.

XX
AC AAA63292:

XX DT 02-FEB-2001 (first entry)

C. trachomatis LGV II clone 12-G3-83 coding sequence.

XX
KW chlamydial infection: sexually transmitted disease;

KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW Chlamydia infection; sexually transmitted disease;

KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;

KW trachoma; blindness; acute respiratory tract disease; heart disease; antibacterial; ss.

XX OS Chlamydia trachomatis.

XXS

PN	WC200034483-A2.	
XX		
PD	15-JUN-2000.	
XX		
PF	08-DEC-1999;	99WO-US029012.
XX		
PR	08-DEC-1998;	98US-00208277.
PR	08-APR-1999;	99US-00288594.
PR	01-OCT-1999;	99US-00410568.
PR	22-OCT-1999;	99US-00426571.
XX		
PA	(CORI-) CORIXA CORP.	
XX		
PI	Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;	
XX		
DR	WPI; 2000-431303/37.	
XX		
PT	Isolated polypeptide for diagnosis and treatment of Chlamydia infection	
PT	comprises immunogenic portion of Chlamydia antigen, which comprises amino	
PT	acid sequence encoded by polynucleotide sequence.	
XX		
PS	Claim 1; Page 128-129; 256pp; English.	

Query Match 8.7% Score 187.2; DB 3; Length 1537;
Best Local Similarity 53.0%; Pred. No. 1.9e-36;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;
SQ Sequence 1537 BP; 459 A; 307 C; 357 G; 414 T; 0 U; 0 Other;

1572 CCATCAATGATGCCCTATGGTAGGGCACGAAATGATGCGACTCGTGATGTGATAAACAAATG 1631

RESULT 13	
AAH56196	
ID	AAH56196 standard; DNA; 1537 BP.
XX	
XX	
AC	AAH56196;
XX	
DT	05-SEP-2001 (first entry)
XX	
DE	Chlamydia trachomatis LGV II DNA clone 12-G3-83.
XX	
KW	Chlamydia; vaccine; infection; fusion protein; antigen;
KW	pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW	acute respiratory tract infection; Cap1; CT529; OMCB;
KW	polymorphic membrane protein; pmp; thiol specific antioxidant; TSA; ds.
XX	
OS	Chlamydia trachomatis.
XX	
FN	WO200140474-A2.
XX	
PD	07-JUN-2001.
XX	
PF	04-DEC-2000; 2000WO-US032919.
XX	
PR	03-DEC-1999; 99US-00454684.
PR	19-APR-2000; 2000US-00556377.
PR	20-JUN-2000; 2000US-00598419.
XX	
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Probst P, Bhatia A, Skeiky YAW, Fling SP, Schollier J;
XX	
DR	WPI; 2001-374831/39.
XX	
PT	Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT	inflammatory disease, trachoma, acute respiratory tract infections,
PT	atherosclerosis and heart disease.
XX	

Example 1; Page 138-139; 295pp; English.

The present nucleotide sequence is provided in a specification relating to compounds and methods for the treatment and diagnosis of chlamydial infection. The compounds provided include polypeptides and fusion proteins comprising immunogenic portions of Chlamydia antigens and DNA sequences encoding such polypeptides. They are useful for vaccinating against chlamydial infection, which causes pelvic inflammatory disease, trachoma, acute respiratory tract infections, atherosclerosis and heart disease

XX Sequence 1537 BP; 459 A; 307 C; 357 G; 414 T; 0 U; 0 Other;

Query Match 8.7%; Score 187.2; DB 4; Length 1537;
 Best Local Similarity 53.0%; Pred. No. 1.9e-36;
 Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY	1167	TTTCCATGCTGTAGATGCTGAAATGAGACCGCTTCATTTTGATCTCTGGGTTC	1226
Db	186	TTTCCITTTGCTTGATGTAGCAATGAATGGCAGCGATTGCATGCAAGTTTTC	245
QY	1227	GTCCAGATGATTCACATGTTCAATACGGAAATCCTGATTTCAAGTCGCCAACAGGAGC	1286
Db	246	GATCTATGATCGAACAAATTAAATGTTAAACAATCTGCAACAGCTAAAGAGCTACAAGCTA	305
QY	1287	TCCGAGCACAGCTAGACGACGGAACCGCCTGGAGATGACAGTGTCTGTGACGCGTGG	1346
Db	306	TGGAGGCTCAGCTGA---CTGCCATGTCAAGTCAACTGTTGTTGGGATGGCAGCTCC	362
QY	1347	CAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGTTAAAGCTGGGCAACAAC-----	1397
Db	363	CAGCGAATACAAAGCAATCAAAAGATGCTCTTCGCAAGCTTTGAACAACCATCAGCAG	422
QY	1398	AGGCGTACTCAATGCTTTAGGACAGATCGTCTTCTGCTGTGTGTGAGCGCAGGAGTTC	1457
Db	423	ATGGTTTAGCTACAGCTATGGGACAAAGTGCGCTTTTGCAGCTGCCAAGTTTGGAGGAGCT	482
QY	1458	CTCCCGCTGCAGCAAGTCTTATAGGTCATCTCTGTAACACAGCTTTTACAGACCTCAAAAT	1517
Db	483	CCGCAGGAACAGCTGGCACTGTCCAGATGAATGTAAACACAGCTTTTACAGACAGCGTTT	542
QY	1518	CTACAGGTTCT-----GATTATATAAACACAGATATCAGCAGGTTATGATGCTTACAAT	1571
Db	543	CTTCGACTTCTCCAGCTCTTATGCAGCAGCATTTCCGATGATATTCTGCTTACAAA	602
QY	1572	CCATCAATGATGCTATGTPAGGCGACGAATGATGCGACTCGTGATGTGATAACAATG	1631
Db	603	CACGTAACTCTTTATATTCGAAAGCAGAAAGCGGCGTGCAGTCAGCTATTAGTCAAACTG	662
QY	1632	TAAGTACCCCGCTCTCACAGATCGTTCTCTAGACACGAACAGAGCTCGAGGCCAG	1691
Db	663	CAAAATCCCGCGTTTCCAGAGCGTTTCTGGTCTCGCATGAAAGTCAAGGACGCAAGT	722
QY	1692	AAAAACAGATCAAGCCCTCGTCTAGGGTGATTTCTGGCAATAGCAGAACTCTTTGGAGATG	1751
Db	723	CAGATGCTAGCCAAAAGACGAGCAGAAACTATTGTCAGAGATAGCCAAAAGTTAGGTGATG	782
QY	1752	TCTATAGTCAAGTTTCGGCACAATCAATCTGTATGATGATATCCAGTGAATCTCTCAAG	1811
Db	783	TATATAGCCGCTTACAGGTTCTGGATTTCTTGATGTCTAGCTGTGAGCAATCCGCAAG	842
QY	1812	CGAATAATCAGGAGATCAGACAAAAGCTTACATCGGCAGTGAACAAAGCTTCCACAGTTTG	1871
Db	843	CAAAATCAGAAGAGATATTGACAGAGCTCAGGCATCTATTAGCAAAGCTTCCAAATTTG	902
QY	1872	GCTATCTTATGTGCAACTTTTCTAATGACTCTACACAGAAGTTTCATAGCTTAATTAGAAA	1931
Db	903	GGTATCCTGCTGTTCAGAATTTCTGTGGATAGCTTTCAGAAAGTTTGTGCACAATTTGAAA	962
QY	1932	GTTTGTTTGTGTAAGATCTAGGACAGAGCTGAATAAAAGCACTTCTCTTTGAAACGA	1991
Db	963	GAGAGTTTGTGATGGGAACCTAGTCTTCGCAAGAACTCTCAAGAGAATGCGTTTAGAAAC	1022

PD 19-JUN-2003.
XX 12-DEC-2002; 2002WO-IB005761.
PF 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX (CHIR-) CHIRON SPA.
PA
XX Grandi G, Ratti G;
XX WPI; 2003-532882/50.
DR P-PSDB; ADD43866.
DR
XX New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.
XX
XX Claim 7; SEQ ID NO 162; 164pp; English.
XX
XX The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This polynucleotide sequence
CC represents one of the 131 DNA sequences encoding the C. trachomatis
CC proteins with immunogenic properties of the invention.
XX
XX Sequence 1944 BP; 612 A; 391 C; 445 G; 496 T; 0 U; 0 Other;
SQ
Query March 8.68; Score 186; DB 9; Length 1944;
Best Local Similarity 48.58; Pred. No. 4.2e-36;
Matches 720; Conservative 0; Mismatches 725; Indels 39; Gaps 6;
606 AGTTGGCGCGCAAAATACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGATTCTTG 665
467 AGTTATTGATAAGTCAATCAGTTAGTTAACTTGAAACACAGATCAGACTTTAAAGG 526
666 ACCTTTAGTTAACTGACTTCCTTCGACCTCTTAAGGCTGCTTCTTCCAACTCTGTAG 725
527 AAACCTTTAAACCAACACAGACTCTGCAGATCAGATTCAGCGGATTAATAGTCAGTTAGAGA 586
726 CAACAAATAACAAAGAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGTCCCAG 785
587 TCACAAATAATCTGCAGATCAAAATTAACAGATCTGGAAGGACAAACATAGTTATG 646
786 GGAATAACGCTGCAATTTGCTCAATCTTTAGTTAGTCAGACAGATGCTTACAGCCACACAGA 845
647 AAGCTGTTCTCACTAAGCGAGGAGAGTTATCAAGCTTCTTGAAGCGGAAATTAAGT 706
846 TAGAGAAAGTGAATTCGATAGGATGATATTTTGCAGACAGACGCTAGTGGAG 905
707 TAGGACAAGCTTTTCAGTCTTAATTTGTGAGTCTGGGATCAAGCCAGGCTGCAGTTCTTC 766
906 CTGTAGAAATGCTAAATCTAATAAGAGTATAGCAACATAGATTAGCTTAAGCGACGA 965
767 AAGCAGCAAAATAATAGCCCAAGTAATATCGACGCCAGAAATAATTAATTAATGATGCTG 826
966 TCCTACTTGTAGACACAAATAGCTGAAGCTTCAGAAAAAGTTCCCGACTCTCCCAATTC 1025
827 CTGAACGAGGTAAACGAGTTAAACACAGAGATACAGGCTAACGGACTCCCTTTAG 886

QY 1026 TTCAGAAGCGGAACAAATGCGTAATACAGGCTCAGAAAGATCTTAAATAATCAAACCTG 1085
Db 887 TGAATAAAGCTGAGGACGATTTAGTCAAGCAAAAGATATTCAGAGATCAAACCT- 945
QY 1086 CAGATGGTTCTGATGTTTCCAAATCCAGGAACACTACAGTTGGAG-----GCTCCAGCAACA 1140
Db 946 --AGTGGTTCCGATATTCCTATCGTTGGTCCGAGTGGTCCAGTCTGCTTCCGAGGAAGTG 1003
QY 1141 AGGAAGTAGTATTGGTAGTAT-----TCGTGTTTCCATGCTCTTTAGATGATG 1187
Db 1004 CGGTAGGAGCGTTGAAATCCTTAAACAATTCAGGAAGAAATTTCTTGTCTTGATGATG 1063
QY 1188 CTGAAAATGAGACCGCTTCCATTTTCATGCTCGGTTTCGTGATGATGATTCACATGTTCA 1247
Db 1064 TAGACATGAATGGCAGCGATTCGAATGCAAGTTCATCTATGATCGAACAATTTA 1123
QY 1248 ATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGAGGCTCGCAGCAACAAGCTAGAGCAG 1307
Db 1124 ATGTAAACAATCTCTGCAACAGCTAAAGAGCTAACAGCTATGGAGGCTCAGCTGA---CTG 1180
QY 1308 CGAAAGCGCTGAGATGACATGCTGCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAG 1367
Db 1181 CGATGTCAGATCAACTGGTTGGTGGGATGGCGAGCTCCAGCGGAAATACAAAGCAATCA 1240
QY 1368 AAGCGCTCTAGGTAAGCTGGCAACAAC-----AGGGCATACTCAATGCTTTAG 1418
Db 1241 AAGATCCTCTTGGCAGCTTTGAAACAACCACTCAACAGATGGTTTAGCTACAGCTATGG 1300
QY 1419 GACAGATCGTTCTGCTGCTGTGTGTGAGCGCAGAGGTTCTCCCGTGCAGCAAGTTCTA 1478
Db 1301 GACAGTGGCTTTTTCAGCTGCCAAGCTTGGAGAGGCTCCGAGGAAACAGCTGGCACCTG 1360
QY 1479 TAGGGTCATCTGTAACAACAGCTTTTCAAGACCTCAAAATCTACAGTTCTGA-----TT 1532
Db 1361 TCAGATGAATGTAACAACAGCTTTTCAAGACAGCGTTTCTTCGACTTCTTCAGGCTCTT 1420
QY 1533 ATAAAAACAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCCATGGTA 1592
Db 1421 ATGCAGCAGCACTTTCGGATGGATATTCGTCTACAAACACTGAACTCTTATATTCGG 1480
QY 1593 GGGCAGAAATGATGCGATCTCGTATGTGATAAACAATGTAAAGTACCCCGCTCTCAAC 1652
Db 1481 AAAGCAAGAACGCGCTGTCAGTCTAGTCACTATTTAGTCAAACTGCAAACTCCGGCTTCCAGAA 1540
QY 1653 GATCCGCTTCTAGACACGACGACAGAAAGCTCSAGGACCAAGAAACAGATCAAGCCCTCG 1712
Db 1541 GCGTTTCTCGTTCTGSCATAGAAAGTCAAGGACGCAAGTGCAGATGCTAGCCAAAGAGCAG 1600
QY 1713 CTAGGGTGAATTTCTGCAATAGCAGAACTCTTTGGAGATGCTCTATAGTCAAGTTTCGGCAC 1772
Db 1601 CAGMAACTATTGTCAGAGATAGCCAAACGTTAGGTGATGTATATAGCCGCTTACAGGTTTC 1660
QY 1773 TACAATCTGATGAGATCTCCAGTCCGATCTCAAGCAATCTCAAGCAATATAGAGGATCAGAC 1832
Db 1661 TGGATTTCTTGTGATGCTACGATTTGTAGCAATCCGCAAGTAAATCAAGAAAGATTTATGC 1720
QY 1833 AAAAGCTTACATCCGCGAGTGACAAAGCTCCACAGTTTCGGCTATCCTTATGTGCACTTTT 1892
Db 1721 AGNAGCTCAGGATCTATTAGCAAGCTCCCAATTTGGGTATCTCTGCTGTTCAAGAT 1780
QY 1893 CTAATGACTCTACACAGAAGTTCTATAGTAAATTAAGAAAGTTTGTGTGCTGAAGGATCTTA 1952
Db 1781 CTGCGGATAGCTTTGCAAGATTTGCTGCGCAATTTGMAAGAGAGTTTGTGTGAGGGAAC 1840
QY 1953 GCACAGCAGCTGAATAAAGACACTTTCCTTTGAACAGCACTTCTGTTGATTCAGCAGG 2012
Db 1841 GTAGTCTCCGCAATCTCGAGAGAATCGGTTTAGAAAAACAGCCCGCTTTCATTCACAGG 1900
QY 2013 TCGTGTCAATTCGGCTCTCTATATTCGTGGTTATCTCCATAA 2056
Db 1901 TGTGTGTAACATGCTTCTCTATTCCTCTGGTTATCTTCTTTAA 1944

Wed Mar 24 14:13:14 2004

us-10-608-559-1.rng

Page 18

Search completed: March 23, 2004, 20:02:57
Job time : 840.202 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:15 ; Search time 8378.96 Seconds
(without alignments)
11152.641 Million cell updates/sec

Title: US-10-608-559-1
Perfect score: 2156
Sequence: 1 ataaatctttaaacaagg.....gatcttcaataaataaaga 2156

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mt.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2156	100.0	2156	6	AX045131 Sequence
2	2156	100.0	10421	1	AE001554 Chlamydia
3	2156	100.0	12548	1	AE002165 Chlamydia
4	2156	100.0	300066	1	AE017159 Chlamydia
5	2156	100.0	300550	1	AP002547 Chlamydia
6	2143.4	99.4	110000	6	Continuation (9 of
7	1956	90.7	1956	6	AX338293 Sequence
8	1956	90.7	1956	6	AX349491 Sequence
9	1956	90.7	1956	6	AX361992 Sequence
10	1848.8	85.8	1852	6	AX045133 Sequence
11	1465.8	68.0	2238	6	AX045137 Sequence
12	1458	67.6	2148	1	CHT76XDA
13	1455	67.5	1456	6	AX045135 Sequence
14	395	18.3	304769	1	AE016994 Chlamydia
15	187.2	8.7	1537	6	BD264444 Compounds
16	187.2	8.7	1537	6	AR229257 Sequence
17	187.2	8.7	1537	6	AR308168 Sequence
18	187.2	8.7	1537	6	AR321875 Sequence
19	187.2	8.7	1537	6	AX155814 Sequence
20	187.2	8.7	1537	6	AX361664 Sequence
21	187	8.7	10954	1	AE001333 Chlamydia
22	186	8.6	1944	6	AX786284 Sequence
23	185.4	8.6	1941	6	AX362019 Sequence
24	184	8.5	1171	6	AX338243 Sequence
25	184	8.5	1834	6	AX338268 Sequence
26	183	8.5	1983	6	AX338258 Sequence
27	170	7.9	1550	6	BD243971 Chlamydia
28	170	7.9	1550	6	AR212731 Sequence
29	170	7.9	1550	6	AR428600 Sequence
30	148.4	6.9	150	6	AR144067 Sequence
31	118	5.5	14698	1	AE002357 Chlamydia
32	55.4	2.6	2000	6	AX655393 Sequence
33	54.4	2.5	7218	6	I66494 Sequence
34	52.4	2.4	110000	2	PFMAL13P2_3
35	52.4	2.4	110000	2	PFMAL13_23
36	51.6	2.4	2205	6	AR317736 Sequence
37	51.4	2.4	154160	2	BX571709
38	50.6	2.3	423	1	MC049
39	49.4	2.3	5098	1	NCU51235
40	49	2.3	1990	6	AR268644 Sequence
41	49	2.3	2085	6	AR008322 Sequence
42	49	2.3	2085	6	AR027306 Sequence
43	49	2.3	2085	6	AR035771 Sequence
44	49	2.3	2085	6	AR037631 Sequence
45	49	2.3	2085	6	AR078816 Sequence

ALIGNMENTS

RESULT 1
AX045131
LOCUS AX045131
DEFINITION Sequence 1 from Patent WO0066739.
ACCESSION AX045131
VERSION AX045131.1 GI:11343730
KEYWORDS
SOURCE Chlamydia pneumoniae
ORGANISM Chlamydia pneumoniae
REFERENCE
1 Mardin, A.D., Omen, R.P., Wang, J. and Dunn, P.
i (Chlamydia) antigens and corresponding dna fragments and uses
thereof
Patent: WO 0066739-A 1 (09-NOV-2000)

AX045131 Sequence 1 from Patent WO0066739.
2156 bp DNA linear PAT 24-NOV-2000

FEATURES		Aventis Pasteur Limited (Ca)	
source	Location/Qualifiers		
CDS	1..2156		
	/organism="Chlamydomytila pneumoniae"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:83558"		
	101..2056		
ORIGIN	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAC17329.1"		
	/db_xref="GI:11343731"		
	/db_xref="REMBL:CA17329"		
	/translation="MVNDIGPGPIDETETPPADLSAOGLEASANKSAEORLAGAE AKPKESKTSVERNSILRSVNALMSLADKLGIASSSSSSRSADVDSITATAPT PPFTDDYKTOAQYAYDIFSTSLADIQALVSLQDAVTWKDTAATDEITAAEW ETKADAYKVGNAQITELAKYASDQALIDSLGKITSFLLORALLOSANNKAEALL KEMQNPVPPKTPAIAOSLVDOTDATTATQDIEKQGNADIRAYFAGONSAGVENAKSN NSINISDAKAAITAKTOIAEAOQKFFPDSILQEAQMV:QAEKDLKNIKPADGSDV PNPGTTGGSKQSSSISIRVSMILLDDAENETASILMSGFRQMHMENTENPDSQA QOELAAQARAALAAAGDDSAALADAOALEALGKAGQOQILNALGOIASAAVVA GVPPAAASSIGSSVKOLYKTSKGSYKTOISAGYDAYKSLINDAYGRANDATRDVI NNVSTPALTRSPRARTPRGPEKTDQALARVISGNSRTLDGVYSQVLSALQVNOITQ SNPQANNEIRKQLISATVPKPGFPYVQLSNDSTQKFIKLSLFAEGSRTAABIK ALSFETSLFIIQQVLNVIGSLSYGLQ"		
	Query Match 100.0%; Score 2156; DB 6; Length 2156;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATAAATCTTTAAACAGCGTCGCATTAAATATTAGTGAGAGCTTTTATTTTTT	60
Db	1	ATAAATCTTTAAACAGCGTCGCATTAAATATTAGTGAGAGCTTTTATTTTTT	60
Qy	61	ATAAATAAATAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	120
Db	61	ATAAATAAATAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	120
Qy	121	AGTCTCTATAGCAAAACAGAACGACACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA	180
Db	121	AGTCTCTATAGCAAAACAGAACGACACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA	180
Qy	181	GCGAGTGCAGCAATAAGATGCGGAAGCTCAAGAGATAGCAGTGGGAAAGCTAAGCC	240
Db	181	GCGAGTGCAGCAATAAGATGCGGAAGCTCAAGAGATAGCAGTGGGAAAGCTAAGCC	240
Qy	241	TAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTGGTTCTGAGTGAATGC	300
Db	241	TAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTGGTTCTGAGTGAATGC	300
Qy	301	TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTTAGTAACAGCTCGTCTTCTACTAG	360
Db	301	TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTTAGTAACAGCTCGTCTTCTACTAG	360
Qy	361	CAGATCTGCACAGTGGACTCAACAGACGACCGCACTAGCCCTCTCCACCCAGTT	420
Db	361	CAGATCTGCACAGTGGACTCAACAGACGACCGCACTAGCCCTCTCCACCCAGTT	420
Qy	421	TGATGATTATTAAGACTCAAGCGCAACAGCTTACGATACACTCTTACCTCAACATCACT	480
Db	421	TGATGATTATTAAGACTCAAGCGCAACAGCTTACGATACACTCTTACCTCAACATCACT	480
Qy	481	AGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACTAATAAAGGATAC	540
Db	481	AGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACTAATAAAGGATAC	540
Qy	541	AGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGGGAAACTAAGAAATGCCGATGC	600
Db	541	AGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGGGAAACTAAGAAATGCCGATGC	600
Qy	601	AGTTAAAGTTGGCGGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGAT	660
Db	601	AGTTAAAGTTGGCGGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGAT	660

Db	601	AGTTAAAGTTGGCGGCAAAATTTACAGAAATAGCGAAATATGCTTCGGATACCAAGCGAT	660
Qy	661	TCCTTGACTCTTTAGGTAAACGTGACTTCCTTGAGACTCTTTACAGGCTGCTCTTCTCCAAATC	720
Db	661	TCCTTGACTCTTTAGGTAAACGTGACTTCCTTGAGACTCTTTACAGGCTGCTCTTCTCCAAATC	720
Qy	721	TGTAGCAAAACATAACAAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATACCCAGTAGT	780
Db	721	TGTAGCAAAACATAACAAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATACCCAGTAGT	780
Qy	781	CCCAGGAAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC	840
Db	781	CCCAGGAAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC	840
Qy	841	ACAGATAGAGAAAGATGGAATGCGATAGGATGCAATATTTTGCAGACAGAAACGCTAG	900
Db	841	ACAGATAGAGAAAGATGGAATGCGATAGGATGCAATATTTTGCAGACAGAAACGCTAG	900
Qy	901	TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTACAGTAAAGC	960
Db	901	TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTACAGTAAAGC	960
Qy	961	AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC	1020
Db	961	AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC	1020
Qy	1021	AATTCTTTCAAGAGCGGAACAAATGGAATACAGGCTGAGAAAGATCTTAAATATCAAA	1080
Db	1021	AATTCTTTCAAGAGCGGAACAAATGGAATACAGGCTGAGAAAGATCTTAAATATCAAA	1080
Qy	1081	ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACATACAGTTGGAGGCTCCAAAGCAACA	1140
Db	1081	ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACATACAGTTGGAGGCTCCAAAGCAACA	1140
Qy	1141	AGGAGTACTATTGGTAGTATTGCTGTTCCATGCTCTTAGATGATGCTGAAATGAGAC	1200
Db	1141	AGGAGTACTATTGGTAGTATTGCTGTTCCATGCTCTTAGATGATGCTGAAATGAGAC	1200
Qy	1201	CGTTTCCATTTTGATGCTGCTGTTTCTGATGATTTACATGTTTCAATACGGAATATCC	1260
Db	1201	CGTTTCCATTTTGATGCTGCTGTTTCTGATGATTTACATGTTTCAATACGGAATATCC	1260
Qy	1261	TGATTTCTCAAGCTGCCCAACAGGAGCTCGCAGCACAACTAGAGCAGCGAAGCCGCTGG	1320
Db	1261	TGATTTCTCAAGCTGCCCAACAGGAGCTCGCAGCACAACTAGAGCAGCGAAGCCGCTGG	1320
Qy	1321	AGATGACAGTCTGCTGCTGAGGCTGGCAGATGCTCAGAAAGCTTTAGAACGGCTCTAGG	1380
Db	1321	AGATGACAGTCTGCTGCTGAGGCTGGCAGATGCTCAGAAAGCTTTAGAACGGCTCTAGG	1380
Qy	1381	TAAAGCTGGGCAACAAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT	1440
Db	1381	TAAAGCTGGGCAACAAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT	1440
Qy	1441	TGTGAGCGCAGGAGTCTCTCCCGCTGCGAAGTTCTATAGGCTCATCTCTATAAACAGCT	1500
Db	1441	TGTGAGCGCAGGAGTCTCTCCCGCTGCGAAGTTCTATAGGCTCATCTCTATAAACAGCT	1500
Qy	1501	TTACAAGACCTCAAAATCTACAGTCTGATTTAATAACACAGATATCAGCAGGTTATGA	1560
Db	1501	TTACAAGACCTCAAAATCTACAGTCTGATTTAATAACACAGATATCAGCAGGTTATGA	1560
Qy	1561	TGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGAAATGATCGCATCTGCTGATGT	1620
Db	1561	TGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGAAATGATCGCATCTGCTGATGT	1620
Qy	1621	GATATAAATAATGTAAGTACCCCGCTCTCACACGATCGGTTCTTAGAGCACGAAACAGAGC	1680
Db	1621	GATATAAATAATGTAAGTACCCCGCTCTCACACGATCGGTTCTTAGAGCACGAAACAGAGC	1680
Qy	1681	TCGAGGACCGAAGAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAAC	1740
Db	1681	TCGAGGACCGAAGAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAAC	1740

QY 1741 TCTTGAGATGCTCTAGTCAAGTTTCGGCACTACAACTCTGAATCAGATCATCCAGTC 1800
Db 1741 TCTTGAGATGCTCTAGTCAAGTTTCGGCACTACAACTCTGAATCAGATCATCCAGTC 1800
QY 1801 GAATCTCAAGCGAATAATGAGAGATCAGACAAAAGCTTACATCGGCGAGTCAAAAGCC 1860
Db 1801 GAATCTCAAGCGAATAATGAGAGATCAGACAAAAGCTTACATCGGCGAGTCAAAAGCC 1860
QY 1861 TCCACAGTTTGGCTATCTCTATGTCAGTCTTCTAATGACTCTACACAGAAGTTCATAGC 1920
Db 1861 TCCACAGTTTGGCTATCTCTATGTCAGTCTTCTAATGACTCTACACAGAAGTTCATAGC 1920
QY 1921 TAAATPAGAAAGTTTGTTCGTAAGGATCTAGGACAGAGCTGAATAAAGCACTTTC 1980
Db 1921 TAAATPAGAAAGTTTGTTCGTAAGGATCTAGGACAGAGCTGAATAAAGCACTTTC 1980
QY 1981 CTTTGAAGCAACTCTCTGTTTATTCAGCAGGTCTGCTCAATATCGGCTCTCTATATTC 2040
Db 1981 CTTTGAAGCAACTCTCTGTTTATTCAGCAGGTCTGCTCAATATCGGCTCTCTATATTC 2040
QY 2041 TGGTTATCTCCATAACAACTCAAGTCTGTTTGGAGAGATTATTAATGCTTTTGT 2100
Db 2041 TGGTTATCTCCATAACAACTCAAGTCTGTTTGGAGAGATTATTAATGCTTTTGT 2100
QY 2101 AAGGCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATPAAATAAAGA 2156
Db 2101 AAGGCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATPAAATAAAGA 2156

RESULT 2
AB001654/c
LOCUS
DEFINI
ACCESS
VERSIC
KEYWOR
SOURCE
ORGAN
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

10421 bp DNA linear BCT 01-DEC-2000
amydia pneumoniae section 70 of 103 of the complete genome.
1654 AB001363
1654.1 GI:4377031
mydophila pneumoniae CWL029
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia. 1 (bases 1 to 10421)
Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R.W., Olinger, L., Grimwood, J., Davis, R.W. and Stephens, R.S.
Comparative Genomes of Chlamydia pneumoniae and C. trachomatis
Nat. Genet. 21 (4), 385-389 (1999)
10192388
2 (bases 1 to 10421)
Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Olinger, L., Grimwood, J., Davis, R.W. and Stephens, R.S.
Direct Submission
Submitted (01-DEC-1998) Program in Infectious Diseases, University of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA
Location/Qualifiers
1. 10421
/organism="Chlamydia pneumoniae CWL029"
/mol_type="genomic DNA"
/strain="CWL029"
/db_xref="taxon:115713"
133..2751
/gene="CpN0727"
133..2751
/gene="CpN0727"
/codon_start=1
/transl_table=11
/product="CpN0727 hypothetical protein"
/protein_id="AAL18866.1"
/db_xref="GI:4377032"
/translation="WSTFISQNRKLTISGESTRIKLDHKYSGFDPSPVAINBELN
SGIYALHMLNLOSENTNVAALNPNFTTFTSWTDYKHSRQASPPSPSSQPT
DIVSAALALVIVIDGGLAELVASVTEIDTGLASTISTVRLMASVGLTITABQEK
VVFSSYVPSEKHLHVQKAEIQAKQBEIKAVLEAKGVSTEEIRAILKEYPDYI
AADFKEFIEEPHLYRAKVGAPIQEMENALQLLTPPATPDVNEVNGNTLSTI

LQADDAIKQAPALGGQDEIITLQTLVPLVDKTTTAKAFDLIYTATQLPNTASLKL
YLTDQIAEYKQKTKVYONISQNLSETRVNNRSMLETLQSLMFOQAOCNCFVWIS
QANALNIAITNKYLSAVLTSMEMVGLLCLSYMYERLADDEKALFDSKNEYLIPIHI
VYGSWNGWIAKMAAYQELAEYSLGTAVTSDQDKAIQTRGNFRHFFHNGD
QMGYPANETVFGNCLTANGAIQPDGGFIREAMTNVGTVEADYVNSQRIILNFNTA
ATAVHQLQIAELQKQADDLDFKASFTENRKFVAAWITSESLGALISMLNSQ
LPQAFALPLIETEFNLANALNSLITNEFTSTVYVSSLYVQSQTQNLQF
AGDYVETLAAAREYIYRDTARCKOALNLVNGKLOKINSLPGATSKQKOBMLNATF
YQYSLSYLNLQTLVLESLAGLWTLQTLTNNKYDKSVKTESFDWIPTLAALESF
LTSFPNLSATGGGLPFTQVSDQQTTSQGTQQLNLQNTTIQEWTLVSTSMQ
VNLGLSLQAGAIYSN"
complement (3078)..5033)
/gene="CpN0728"
complement (3078)..5033)
/gene="CpN0728"
/codon_start=1
/transl_table=11
/product="CHLPN 76 kDa Homolog_1 (CT622)"
/protein_id="AAL18867.1"
/db_xref="GI:4377033"
/translation="WNPFGPIDETERTPPADLSAOGLEASRANKSAEQRLAGAE
AKPKESKDSVERWSILRSVNAVMSLADKLGSIASSNSSSTSRSDVDSITATPTP
PPFDDYKTAQATYDTITFTSLADIQAALVSLQDRAVTNIKOTAATDEETATAEW
ETKNADA VKVQAITEAKYASDNQALDLSGLKLTSLDQLQAALQSVNNAKAEEL
KEMQNPVPGKTPAIQSLVDQDATATQIEKDNATRDAYFAGONASGAVNAKSN
NSINIDSAKAIATQIAEAKQPPDPILOEAEQMTIOAEKDLKNIKPADQSDV
PNPTVGSKQGGSSIGSIRVSLMLDDAENETASILMSGFQMTLHMENTENPOQAA
QOELAAQARAAGADDSAAALADAQKALEAALGKAGQOQGLNALQGLASAAVSA
GVPPAASSIGSSVKLYKTSKTSYDTQISAGYDAYKSDINADYARNDATRDVI
NNVSTPALTRSPRAREPEKTDQALARVISGNSRTLDGVSQVSAQSVMQIIQ
SNPQANNEEIRKLTSAVTPEPFGYVQVQLSNDSTQKFIARLESPLAEGSRTAAEIK
ALSFTNLSFTQVVLNIGSLYGLQ"
complement (5145)..6413)
/gene="CpN0729"
complement (5145)..6413)
/gene="CpN0729"
/codon_start=1
/transl_table=11
/product="CHLPN 76 kDa Homolog_2 (CT623)"
/protein_id="AAL18868.1"
/db_xref="GI:4377034"
/translation="MKQVYQWLASVLLALITISYAEPLSEQKVKSHYTTTLDVK
DYLKRGFVETRKQDGLRIAGDVARWLYFREDIKPDKDKKYNPLPVRNYSFYL
YIDRAERNWLSKMNMTAIGAGNTAGVDINRAFLGYFKNPETDPTDFMEIGRS
GLGLDFSEVOFQSNFQDGLHLYWTRLSKDPYQVIVHGGFVNVNNTKKHAWVEGI
LNRLLPKOPFVKCSVDWNTFVPSSTTTEKATNAMYKYCVWQVLGKHSQVPVING
OKPLYLNGALMPLAKATKTTLNGKENLAWFGLGLGKRGKADGASATVYEVGA
LSVPEIDVSGIRGNLLKFWFAQALAAANYDPEANGFTNYKGFSAIYMGITDLSFR
AIGAYSPANDKLGSDFTFRKFDGLISAF"
complement (6516)..8159)
/gene="mvIN"
/note="synonym: CpN0730"
complement (6516)..8159)
/gene="mvIN"
/codon_start=1
/transl_table=11
/product="Integral Membrane Protein"
/protein_id="AAL18869.1"
/db_xref="GI:4377035"
/translation="MSRKDNVSLARSIPNLSGTFCSRITGIFREIMATYFGADPI
VAFWLGRFTVFRKQYGLLEQAFIPHFEPFRAQSLDMRAAFFFFRFRSLKLGSTI
IFTLLEAVLMVLQYVEEGYDTMLTILLPCGIFIMMYNANGALHLCNKKFPGVG
LAPVNVNIWIFFVIAARHSDPRERIIGLSVALVYFFFWLITVPGWKFLEAKSP
POEHSVRALLAPLSGLITSSIFQLNLSIDICLARYVHEIGPLVMSLYIOLPIH
LFGFGFTVLLPAISRCVQREDHRLKMKFVLTLMSVMIINTAGYLIALPGVRV
LYEHGLPQSAVYAVRULRGYASIIIPMALAPLVULFYAQRYAVPLFIFIGITALA
NIVLSLMGRWLKDVSGISATISATVAVYFLWYSSKRLPMYKLSLWESIRRSIK
VGFMTLACMTLTGLNLTQTYVIFLNPPLPLAWPLSSITAQAFLESCEIFLAFI
FGFAKLREVDLNLASFEYWRGQGLLQROHVNQDTQN"
8143..8319
/gene="CpN0731"
8143..8319
/gene="CpN0731"

gene	8654..9535	/codon_start=1	
	8654..9535	/transl_table=11	
	8654..9535	/product="hypothetical protein"	
	8654..9535	/protein_id="AADI8870.1"	
	8654..9535	/db_xref="GI:4377036"	
	8654..9535	/translation="MSFLLLKPSPPPLKQDYLFDISPYTSSSEITTTGGSYFKLNKASL	
	8654..9535	QSSTRLRSISIIIS"	
	8654..9535	/gene="info"	
	8654..9535	/note="synonym: CPn0732"	
	8654..9535	/gene="info"	
	8654..9535	/codon_start=1	
	8654..9535	/transl_table=11	
	8654..9535	/product="Endonuclease IV"	
	8654..9535	/protein_id="AADI8871.1"	
	8654..9535	/db_xref="GI:4377037"	
	8654..9535	/translation="MKVLPPSPILLGAHTSTAGGLKNALYEGRDYGASTVQIFPTANQ	
	8654..9535	RQWRALKEEVIEDFAALKKEIDLSITMSHAGYLINPGAPDVLLEKSRIGIYQIBL	
	8654..9535	DCITLIGSPNFPHPGALKSKSDCKMKIVSFSQAPLFDSPSPVLVLETTAGQGT	
	8654..9535	LIGNSEEEGLYQNLKQNIPIGVCVDTHIPAAGYDITSPOGWEDVLNEFDEYVGLS	
	8654..9535	YLRAFLINDSMFPLGANKORHAPLGEYGYIKSEFKFLMTDERTRKIPKYLETPGGDEN	
	8654..9535	WQKEIGELLKFKSKURDS"	
	8654..9535	/complement(9654..10283)	
	8654..9535	/gene="rs4"	
	8654..9535	/note="synonym: CPn0733"	
	8654..9535	/complement(9654..10283)	
	8654..9535	/gene="rs4"	
	8654..9535	/codon_start=1	
	8654..9535	/transl_table=11	
	8654..9535	/product="S4 Ribosomal Protein"	
	8654..9535	/protein_id="AADI8872.1"	
	8654..9535	/db_xref="GI:4377038"	
	8654..9535	/translation="MARYCGPKNRVARRFGANIFGRSNPLLKPHPPGQHGOMORKKK	
	8654..9535	SDYGLQLEKQKLCACYGKMEKLVKAFKEVHKQNVAKOMFLERFECRLNMYVRM	
	8654..9535	FGAKTTFAAQQLVAHGHLVNGPRVDRSPFLRPGMQVLSLKEKSLQSVKDALRSKD	
	8654..9535	ESLSFYISLDKTKFGFGEUUVSEQDCEIAQPLPINISVVCEFSHRT"	
ORIGIN			
Query Match	100.0%;	Score 2156;	DB 1; Length 10421;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2156;	Conservative	0; Mismatches	0; Indels 0; Gaps 0;
QY	1	ATAAAATCTTTAAAAACAGGCTCGCATTAATTAATTAAGTCAGAGCTTTTTTTTATTTTT	60
Db	5133	ATAAAATCTTTAAAAACAGGCTCGCATTAATTAATTAAGTCAGAGCTTTTTTTTATTTTT	5074
QY	61	ATAATAAACTAAAGATTTTTTATTTTTTTCAGTTTTTATNGTAAATCCCTATTGGTCC	120
Db	5073	ATAATAAACTAAAGATTTTTTATTTTTTTCAGTTTTTATNGTAAATCCCTATTGGTCC	5014
QY	121	AGTCTCTATAGACGAAACAGACACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA	180
Db	5013	AGGTCCTATAGACGAAACAGACACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA	4954
QY	181	GGCGAGTCAGCAATAAAGAGTCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC	240
Db	4953	GGCGAGTCAGCAATAAAGAGTCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC	4894
QY	241	TAAAGAATCTAAGACCGATTCTGTAGACGATGGAGCATCTTTCGCTCTGCGAGTGAATGC	300
Db	4893	TAAAGAATCTAAGACCGATTCTGTAGACGATGGAGCATCTTTCGCTCTGCGAGTGAATGC	4834
QY	301	TCTCATGAGTCTGGCAGATAAAGTCGGATTGCTTCTAGTAACAGCTCGTCTTCTACTAG	360
Db	4833	TCTCATGAGTCTGGCAGATAAAGTCGGATTGCTTCTAGTAACAGCTCGTCTTCTACTAG	4774
QY	361	CAGATCTCAGAGCTGGATCTAACGACAGCAGCGCACTTACGCTCTCTCCACCAACGTT	420
Db	4773	CAGATCTCAGAGCTGGATCTAACGACAGCAGCGCACTTACGCTCTCTCCACCAACGTT	4714
QY	421	TGATGATTATAGACTCAGCGCGCAACAGCTTACGATACTATCTTTTAACTTCAACATCACT	480

FEATURES	source
Location/Qualifiers	
1. .12548	
/organism="Chlamydomophila pneumoniae AR39"	
/mol_type="genomic DNA"	
/strain="AR39"	
/db_xref="taxon:115711"	
/note="synonym: Chlamydia pneumoniae AR39"	
297. .965	
/gene="CP0011"	
297. .965	
/gene="CP0011"	
/note="similar to SP:P31218 GB:X71492 PID:296947 GB:U00096 PID:1736770; identified by sequence similarity; putative"	
/codon_start=1	
/transl_table=11	
/product="uridine kinase"	
/protein_id="AAF37907.1"	
/db_xref="GI:7188950"	
/translation="MLMMLMIIGITGGGAGKTTLTQNIKEIFGEDVSVICODNVYK DRSHYTPERANLIWDHDAFDNDLIJSDIKELKNEIVQAPVDFVLGNRSKTRLET IYPSXIVLVEGILVFENQELRDLMDIRIFVDTDADERILRMVRDVQCGDSVDCIMS VNSK"	
complement(1047. .2021)	
/gene="CP0012"	
complement(1047. .2021)	
/gene="CP0012"	
/note="conserved hypothetical protein; identified by Glimmer2; putative"	
/codon_start=1	
/transl_table=11	
/product="conserved hypothetical protein"	
/protein_id="AAF37908.1"	
/db_xref="GI:7188951"	
/translation="WEKKYALAYVITVDNPHEEIALHKFFLEDLDVSCRIVTSE GNGQFSQVEHAELYQWLKERNPFSKIKFHHIKENIFRITVKYRKELALGCE LDLSQAKHISQWHEKQLKQNCUILDRNNYEWKIGHFDNATLPDIQTFRFFPEYA EKLAQBCDPTTPYVMCTGIRCELYSPVILEKGFKEVYQLDGGVIAQQOQVGTGW LGKLFIDPDAIDPDAIECDPAIECCHQCTPDAVNCANTCNALFELCCDCBII OHQCCFFECDSRNVKFDNSRGNKDPBRAHICETISNPSASCTT"	
gene	
CDS	

```

CDS
/gene="CP0013"
/238..2867
/note="CP0013"
/note="similar to SP_P81288; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein S4"
/protein_id="AAF37909.1"
/db_xref="GI:7188952"
/translation="MARYCGKNAVRRFGANIIGRGNPILLKKPHPPGQHMORKKY
SDYGLQEKKQKLKACYGMLMEKLVKAFKVIHKQGNVAKMFLERECRLDNVYEM
GPAKTIAAQQLVAGHILNIGRVDRDSFFLRFGMKSLAEKSRQSGVADDALESKD
ESSSPISLDKTKFGKGLVSPQDQQLAQPLPINISVCFELSHRT"
complement(2986..3867)
/gene="CP0014"
/complement(2986..3867)
/note="CP0014"
/note="similar to SP_P12638 GB_M22591 PID:146954
PID:405898 GB_U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="endonuclease IV"
/protein_id="AAF37910.1"
/db_xref="GI:7188953"
/translation="MKVLPPPSILLGAHTSTAGGLKNAIYEGRDIGASTVQITFANO
RQWRKALKEVIDFKAALKEITLDSYIMSHAGLYINGPADPVLKSRIGIYOEIL
DCITGLTSFVNFHAKSKEDCMNKIVSSFQSAPLDSSPVLVLLTETAGQGT
LIGNSFELGVLQNLKNQPIGVCDVTCHIFAAQYDITSPQGWEDVLNDEFYVGLS
YLAFLHLDNMFPLGAKNDKRAHPLGEGYIGKESFKLTMDERTKIPKYLTPGCPEN
WKEIGELLLKSKNRDS"

```


661 TCTTGACTCTTTAGGTAAGTCTCTTCTGACCTCTTACAGGCTGCTCTTCTCAATC 720
Db
8048 TCTTGACTCTTTAGGTAAGTCTCTTCTGACCTCTTACAGGCTGCTCTTCTCAATC 8107
Qy 721 TGTAGCAACAATTAACAAGAGCTGAGCTTCTTAAAGAGATGGAAGATTAACCCAGTAGT 780
Db 8108 TGTAGCAACAATTAACAAGAGCTGAGCTTCTTAAAGAGATGGAAGATTAACCCAGTAGT 8167
Qy 781 CCCAGGGAAGCCCTGCTCAATTTGCTCAATTTTGTAGTTGATCAGACAGATGCTACAGCGAC 840
Db 8168 CCCAGGGAAGCCCTGCTCAATTTGCTCAATTTTGTAGTTGATCAGACAGATGCTACAGCGAC 8227
Qy 841 ACAGATAGAGAAAGATGGAAGATGCGAATAGGATGATATTTTGCAGGACAGAACGCTAG 900
Db 8228 ACAGATAGAGAAAGATGGAAGATGCGAATAGGATGATATTTTGCAGGACAGAACGCTAG 8287
Qy 901 TGGAGCTGTAGAAATGCTTAATCTTAATTAACAGATTAAGCAACATAGATTGAGTAAAGC 960
Db 8288 TGGAGCTGTAGAAATGCTTAATCTTAATTAACAGATTAAGCAACATAGATTGAGTAAAGC 8347
Qy 961 AGCAATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 1020
Db 8348 AGCAATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 8407
Qy 1021 AATTTCTCAAGAGCGGAACAATAGTGAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
Db 8408 AATTTCTCAAGAGCGGAACAATAGTGAATACAGGCTGAGAAAGATCTTAAAAATATCAA 8467
Qy 1081 ACCTGCAGATGCTGTGATGTTCCAAATCCAGAACTACAGTTGAGGCTCCAGCAACA 1140
Db 8468 ACCTGCAGATGCTGTGATGTTCCAAATCCAGAACTACAGTTGAGGCTCCAGCAACA 8527
Qy 1141 AGGAAGTAGTATTGCTGATGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 1200
Db 8528 AGGAAGTAGTATTGCTGATGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 8587
Qy 1201 CGCTTCCATTTTGTGCTGGTTCGTCAGATGATCAATGTTTCAATACGAAAAATCC 1260
Db 8588 CGCTTCCATTTTGTGCTGGTTCGTCAGATGATCAATGTTTCAATACGAAAAATCC 8647
Qy 1261 TGATTTCAAGCTGCCCAAGAGCTCGCAGCACAAGCTAGAGCAGAGAAAGCCGCTGG 1320
Db 8648 TGATTTCAAGCTGCCCAAGAGCTCGCAGCACAAGCTAGAGCAGAGAAAGCCGCTGG 8707
Qy 1321 AGATGACAGTGTCTGTCAGCGCTGCGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
Db 8708 AGATGACAGTGTCTGTCAGCGCTGCGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 8767
Qy 1381 TAAAGCTGGGCAACAACAGGGGATCAATGCTTTAGACAGATGCTTCTGCTGTGT 1440
Db 8768 TAAAGCTGGGCAACAACAGGGGATCAATGCTTTAGACAGATGCTTCTGCTGTGT 8827
Qy 1441 TGTAGCGCAGAGTCTCTCCGCTGCGCAGAGTCTATAGGTCATCTGTAAACAGCT 1500
Db 8828 TGTAGCGCAGAGTCTCTCCGCTGCGCAGAGTCTATAGGTCATCTGTAAACAGCT 8887
Qy 1501 TTAACAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGCAGGTTATGA 1560
Db 8888 TTAACAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGCAGGTTATGA 8947
Qy 1561 TGTCTTCAAAATCCATCAATGATGCTTGTAGGACAGAAATGATGCGATCGTGATGT 1620
Db 8948 TGTCTTCAAAATCCATCAATGATGCTTGTAGGACAGAAATGATGCGATCGTGATGT 9007
Qy 1621 GATTAACCAATGTAGTACCCCGCTCTCACAGATCCGTTCTAGAGCAGCAACAAGAGC 1680
Db 9008 GATTAACCAATGTAGTACCCCGCTCTCACAGATCCGTTCTAGAGCAGCAACAAGAGC 9067
Qy 1681 TCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAAC 1740
Db 9068 TCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAAC 9127

1741 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGAGATCATCCAGTC 1800
Db 9128 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGAGATCATCCAGTC 9187
Qy 1801 GAATCTCAAGCGAATAATAGGAGATCAGACAAAGCTTTCATCGGCGAGTGACAAAGCC 1860
Db 9188 GAATCTCAAGCGAATAATAGGAGATCAGACAAAGCTTTCATCGGCGAGTGACAAAGCC 9247
Qy 1861 TCCACAGTTTGGCTATCCTTATGTGCAACTTTCATGACTCTACACAGAGTTTCATAGC 1920
Db 9248 TCCACAGTTTGGCTATCCTTATGTGCAACTTTCATGACTCTACACAGAGTTTCATAGC 9307
Qy 1921 TAAATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGCTGAAATATAAGACACTTTC 1980
Db 9308 TAAATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGCTGAAATATAAGACACTTTC 9367
Qy 1981 CTTTGAACGAACCTCTTGTGTTTATTCAGACAGGCTGCTCAATATCGGCTCTCTATATTC 2040
Db 9368 CTTTGAACGAACCTCTTGTGTTTATTCAGACAGGCTGCTCAATATCGGCTCTCTATATTC 9427
Qy 2041 TGGTTATCTCAATAACCAACACTTAAGTGTTCGTTGGAGAGATTTATGCTTTCGT 2100
Db 9428 TGGTTATCTCAATAACCAACACTTAAGTGTTCGTTGGAGAGATTTATGCTTTCGT 9487
Qy 2101 AAGGCTTGTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 2156
Db 9488 AAGGCTTGTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 9543

RESULT 4
AE017159/c
LOCUS
DEFINITION Chlamydomophila pneumoniae TW-183, section 3 of 4 of the complete genome.
ACCESSION AE017159 AE009440
VERSION AE017159.1 GI:33236383
KEYWORDS
SOURCE
ORGANISM Chlamydomophila pneumoniae TW-183
REFERENCE Chlamydomophila pneumoniae TW-183
AUTHORS Bacteria; Chlamydiae; Chlamydiaceae; Chlamydiales; Chlamydia. 1 (bases 1 to 300066)
Geng, M.M., Schumacher, A., Muehldorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T., Essig, A., Marre, R. and Melchers, K.
The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis
Unpublished
JOURNAL 2 (bases 1 to 300066)
REFERENCE Geng, M.M., Schumacher, A., Muehldorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T. and Melchers, K.
Direct Submission
JOURNAL Submitted (28-MAY-2002) RDR/IT, RDR/FG, RDR/BT, ALTANA Pharma, Byk-Gulden-Str. 2, Konstanz 78467, Germany
FEATURES
source
1. 300066
/organism="Chlamydomophila pneumoniae TW-183"
/mol_type="genomic DNA"
/strain="TW-183"
/db_xref="taxon:182082"
/note="type strain"
biovar: TWAR
126..797
/locus_tag="CbB0543"
126..797
/locus_tag="CbB0543"
/notes="Similar to hypothetical protein CT433 - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF98472.1"
/db_xref="GI:33236384"


```

Db      212781  TAAATTAGAAAGTTGTTGCTGAAGATCTAGACAGCAGCTGAATAAAGCACATTC 212722
QY      1981    CTTTGAACGAACCTCTGTTTATTCAGCAGTCTGCTCAATCGGCTCTCTATATTC 2040
Db      212721  CTTTGAACGAACCTCTGTTTATTCAGCAGTCTGCTCAATCGGCTCTCTATATTC 212662
QY      2041  TGGTTATCTCCAAATACCAACACCTTAAGTGTTCGTTTGGAGAGATTATATGTCGTTGGT 2100
Db      212661  TGGTTATCTCCAAATACCAACACCTTAAGTGTTCGTTTGGAGAGATTATATGTCGTTGGT 212602
QY      2101  AAGCCCTTGTGAGGCTTACCAACACCTAGACACTAGACGATCTTCAATAAATAAGA 2156
Db      212601  AAGCCCTTGTGAGGCTTACCAACACCTAGACGATCTTCAATAAATAAGA 212546

RESULT 5
LOCUS   AP002547/c
DEFINITION Chlamydothila pneumoniae J138 genomic DNA, complete sequence, section 3/4.
ACCESSION AP002547 AB033786 AB033787 AB033816 AB033817 BA000008
VERSION   AP002547.2 GI:10176693
KEYWORDS
SOURCE   Chlamydothila pneumoniae J138
ORGANISM Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydothila.
REFERENCE
AUTHORS Shirai, M., Hirakawa, H., Ouchi, K., Tabuchi, M., Kishi, F., Kimoto, M., Takeuchi, A., Nishida, J., Shibata, K., Fujinaga, R., Yoneda, H., Matsushima, H., Tanaka, C., Furukawa, S., Miura, K., Nakazawa, A., Ishii, K., Shiota, T., Hattori, M., Kuhara, S. and Nakazawa, T.
TITLE    Comparison of outer membrane protein genes omp and pmp in the whole genome sequences of Chlamydia pneumoniae isolates from Japan and the United States
JOURNAL  J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)
MEDLINE  20298986
PUBMED   10839753
REFERENCE
AUTHORS Shirai, M., Hirakawa, H., Kimoto, M., Tabuchi, M., Kishi, F., Ouchi, K., Shiba, T., Ishii, K., Hattori, M., Kuhara, S. and Nakazawa, T.
TITLE    Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWJ029 from USA
JOURNAL  Nucleic Acids Res. 28 (12), 2311-2314 (2000)
MEDLINE  20330349
PUBMED   10871362
REFERENCE
AUTHORS Shirai, M.
TITLE    Direct Submission
JOURNAL  Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
SCHOOL   School of Medicine, Department of Microbiology; 1-1-1
MINAMIKOGUHI, Ube, Yamaguchi 755-8505, Japan
(E-mail: mshirai@po.cc.yamaguchi-u.ac.jp, Tel: 81-836-22-2227, Fax: 81-836-22-2415)
COMMENT  On or before Sep 15, 2000 this sequence version replaced
gi:6172298, gi:6172300, gi:6172396, gi:6172398, gi:8978889.
FEATURES
source   Location/Qualifiers
1..300550
/organism="Chlamydothila pneumoniae J138"
/mol_type="genomic DNA"
/strain="J138"
/db_xref="taxon:138677"
/notes="synonym: Chlamydia pneumoniae (strain J138)"
complement(21..794)
/genes="dapF"
complement(21..794)
/genes="dapF"
/codon_start=1
/transl_table=11
/product="diaminopimelate epimerase"
/protein_id="BAA98725.1"
/db_xref="GI:8978890"
/translation="MAFYSPSTISKFYISYGAGNRFLLGETLPEVEDVFLCQETRVDD
GFLYKPSSCADAQLIIFNSDGRPTMCGNGLRCATAHLASQKGSDDISVSTDSGLYS
GFYSDRVLVRLTDLARASVRESRDPDLPKVEVCVCHTGVPHAVWILPELSTLDL
SILGFLRVHQTSPDGVNPNFVLHGLCQLRVYTVRGEVGETAAGTACGTGALASLVV
SNSYGMKESIQHTHMGELMTYQNRGRVYLGQSVTRDL"
/gene="clpP_1"
complement(763..1338)
/gene="clpP_1"
complement(763..1338)
/codon_start=1
/transl_table=11
/product="CLP protease"
/protein_id="BAA98726.1"
/db_xref="GI:8978891"
/translation="WADGEVHLRDLIELLEEARVFESEPVTEKASDAIKKLWYL
ELKPKPIFVINSFGSDAGFAVMDQIKMTSPVITVTGLAASMGSVLSLCAAP
GRFATPHSRIMHQFSIGGPTGQATDLDIHAREILTKARIIDVYVBEATNQPRDI
EKAIIDRMWMTANEAKDFGLDGLIFSFNDL"
/gene="glyA"
complement(1358..2917)
/gene="glyA"
complement(1358..2917)
/codon_start=1
/transl_table=11
/product="serine hydroxymethyltransferase"
/protein_id="BAA98727.1"
/db_xref="GI:8978892"
/translation="MLKVFKEKFAIVEIFTKVAVVSVLSLHKFLENASGKKGQSLAS
TAYLAADHLNAPFSIGERIIDEKRSQSHLMIASENSSLSVOLAMGNLLTKYC
ESGPAPRYSCENVDIAIEWECVETAKELFAADCAQVPHSGADANLAWAILTHKV
QCFYKLGTYKLTNELTEETLLKAEMSSCVGLGFLNSGHLTHGNVLNVMGKLM
RCFYPDVNPDTCTFYAIEISRLAKYKPKVLIAGYSYSSRLNFAVLKQIADCCSVL
YDMDAFLAGVAGVFEDEENPIYADIVTTTHTKLGRPGGLVLAIREYESTLKA
CPLMAGPHVIAAKTVALKALSVDEKKAHQVNNARLAEFLSHGLRLTGGT
DNHNVVIDLSGLSIGKAEADLSVGIANNVSNLPSDAIGKWDTSGLRLGTPALTTL
GMGIDMEVADIIYKVLNRLSCHVEGSSKKNGELPEAIAQEARDRVRLNLLRFP
IYPEIDLEALV"
/gene="CPj0522"
3107..3778
/gene="CPj0522"
3107..3778
/gene="CPj0522"
3107..3778
/codon_start=1
/transl_table=11
/product="CT433 hypothetical protein"
/protein_id="BAA98728.1"
/db_xref="GI:8978893"
/translation="WTLVLGNQTKARKYQAHYLBILTFPYAKSTPNKRALQFLPQ
ATHVLTSPSTHLFSLRMTSLLSKATIKTKYLCIGESTKERLSFLGQKVVVATQ
EIAEGIFPLQALPSSARILYPHSLARPVIREFLNRTFFSYPHYTVKPKLKNL
LSKYKKIIFTSPTVRAFAKIFRPPEKTYWCQGRMTLQEFQKFSQKQSVLLETGK
SRTSP"
3843..4175
/gene="CPj0523"
3843..4175
/gene="CPj0523"
3843..4175
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98729.1"
/db_xref="GI:8978894"
/translation="MASSATGFGDGTAPSLFPFPATPRYNFKLAFVTLAIVMAL
IATTAIAGLCIHLPLCSFIFLTAIPLYFISRYICSHYARNVYIALDVVDHSLQDMRS
HSPFSDR"
4223..5302
/gene="CPj0524"
4223..5302
/gene="CPj0524"
4223..5302
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98730.1"
/db_xref="GI:8978895"
/translation="MSGFSPSTESSQVSLSVPRDKETAPKQFTIAKIATLIALSL
ALGAVAGISLITVLGNPNFLALLITLALFSWTFVLVYHQTWSKVSNNQKVLQNFQ

```

```
PLKAWOEKNVDCVSNEMQFNNHLPKFKVAIQTDASOPFOPTFELTGLVIEKQST
GIIPNVGPNLINDVATNLSTILYSFLKCKSVWDCKQREGGPAKGEDFFSTFVRL
VKLPNALDQTNLNLSSAKKILFTFLGHVCPKSELPNQOEYIRQALLAYENCL
KAAIESHALVALPFLTSVVEPPEEILPKFEGFYWDNQTPQAFCKRALDLAIQNTALR
YPORSLVILQDFFNTIESQSRSEE"
gene 5642..6406
CDS /gene="CP10525"
5642..6406
/gene="CP10525"
/codon_start=1
/transl_table=11
/product="C7398 hypothetical protein"
/protein_id="BAA98731.1"
/db_xref="GI:8978896"
/translacion="MDALLSILAIQELDIKMLRVRVKEHOKELAKVQSLKSDIRR
KVQEKLEMLNLTQIRDGENRIQEISEQINKLENOQAAVKMDENALQTQENTANK
ERSLHQDLMQKQAGEDLIYSKSLASTENSSVIEKEIFESIKKINEEGKAL
LEQRTBELKATNPELLISYIERLLNNKDRVVPENRVCVSGCHIVLTPQHENVLRKDD
RLIFCHESRILYQBSQVNAQENSTAKRRRRAAV"
complement(6827..7816)
gene /gene="kpsF"
CDS complement(6827..7816)
/gene="kpsF"
/codon_start=1
/transl_table=11
/product="GutQ/KpsF family sugar-P isomerase"
/protein_id="BAA98732.1"
/db_xref="GI:8978897"
/translacion="MSPSMISTDVQDILKQKEAVDFPQFPQPKEMQLAELKILGH
SGWFFSGVKGSCVARKIVATLQSLSERALFFSPVDLHGLVSPGDIVCLFSKS
GETQELLDTVPHLKSRALIVALTSMPSNLAAISDLVILPVAELDPENLIPNTSG
TCOMIFGDFLAMLHPSRGVSLSTYKGNHPSGOVGMKANGKVDNFKPKEVFPCHLG
DKVYSFLEVSFAYGCGVCIVDPQFRLMGIETFDGDLRRSLASVGGVLSLEKMTA
NPKCITSDSDIALQIMRESSPVAVLPVLVDNEENRHVTLGLLHMHTLAKAGLL"
complement(7843..9024)
gene /gene="sucB_2"
CDS complement(7843..9024)
/gene="sucB_2"
/codon_start=1
/transl_table=11
/product="dihydrolipoamide succinyltransferase"
/protein_id="BAA98733.1"
/db_xref="GI:8978898"
/translacion="MIFEPFPKIGTSSGGSVRLKMLGDHVARDEPLIEVSTDKI
ATELPSPKAGRLVFCVNEGDEVASGVLGLLELEBISEADDESTSCPILTSCTKSEA
GSSSSVMTSPSPVLSLAQREGICLDNLQKTACTGKGRVTRQDLEAYISEQVSIPE
IPQGVNRPMSPLRPAIASSLSKSDVPHASLVVDVDDVTDMLNLSIGERQRFLDTH
GVKLITTSIVQCLAQTLKQFPPLNGSLDGTIVMKKSVNVGVAVNLNKEGVVVPVIH
NODRGLVSIKADLSLRANKLPDSEVDQSVTVTFNGMTGALI GMPRIIRYPEV
AILGIGTIQKRVVVRDDDSLAIKRMVYVTLTFDHRVLDLGTGYGSEFLTSLKNRESVTM
G"
complement(9038..10282)
gene /gene="gltT"
CDS complement(9038..10282)
/gene="gltT"
/codon_start=1
/transl_table=11
/product="glutamate symport"
/protein_id="BAA98734.1"
/db_xref="GI:8978899"
/translacion="MKLWKKIFIGLVFVGTGLVLVEDKAIFFKPIGDIFFNLNLSMVVY
PIVFCSMVLGIASIDMKKLGIGIKSVGLYIGLTALAIIVIGLCPAMIFSPNGCDFA
QAQMDSAVTVIDSNKTAAYFLSIIAQVFPSPNVPVSFARGNIIQIIIFALFILRLR
SEGRGRPVRFIDGFEIMLRVMNMTSPAPYGVGASMAWISGNHGLVWLQWGFLL
AYLACLCHATLVFGLVRFPCMSKFSKLSMMDLASCASVASTASSATLPVTRCVS
KMLGSVAEVSQFVLPGATVNMNGTAIFQGMAAVFIQAQYNCPLSLSSLLLVATFAT
SIVGSAVPGGMITLGSVLASVGLPIQIAILAGIDRLDIRDVIPTPMNITLGDVAVTAT
VASGEGLSPYESIKQESVET"
complement(10282..11379)
gene /gene="ycaH"
CDS complement(10282..11379)
/gene="ycaH"
/translacion="MCLWKKIFIGLVFVGTGLVLVEDKAIFFKPIGDIFFNLNLSMVVY
PIVFCSMVLGIASIDMKKLGIGIKSVGLYIGLTALAIIVIGLCPAMIFSPNGCDFA
QAQMDSAVTVIDSNKTAAYFLSIIAQVFPSPNVPVSFARGNIIQIIIFALFILRLR
SEGRGRPVRFIDGFEIMLRVMNMTSPAPYGVGASMAWISGNHGLVWLQWGFLL
AYLACLCHATLVFGLVRFPCMSKFSKLSMMDLASCASVASTASSATLPVTRCVS
KMLGSVAEVSQFVLPGATVNMNGTAIFQGMAAVFIQAQYNCPLSLSSLLLVATFAT
SIVGSAVPGGMITLGSVLASVGLPIQIAILAGIDRLDIRDVIPTPMNITLGDVAVTAT
VASGEGLSPYESIKQESVET"
complement(10282..11379)
gene /gene="ycaH"
CDS complement(10282..11379)
/gene="ycaH"
```

```
Query Match 100.0%; Score 2156; DB 1; Length 300550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTTAAAAACAGGCTCGCATTAATATTATTAGTGAGAGCTTTTATTTTATTTT 60
DB 217698 ATAAATCTTTTAAAAACAGGCTCGCATTAATATTATTAGTGAGAGCTTTTATTTTATTTT 217639

QY 61 ATAATAAACTAAAGATTTTATTATTTTGTAGTTTTTATGTTTAACTTATTTGTTCC 120
DB 217638 ATAATAAACTAAAGATTTTATTATTTTGTAGTTTTTATGTTTAACTTATTTGTTCC 217579

QY 121 AGTCTCTATAGCAAGCAAGACGACCTCCGCGAGATCTTCTGCTCAAGGATTGGA 180
DB 217578 AGTCTCTATAGCAAGCAAGACGACCTCCGCGAGATCTTCTGCTCAAGGATTGGA 217519

QY 181 GCGAGTGCAGCAATAAAGAGTGCAGAGCTCAAGAAATAGCAGGTGCGGAAGCTAAGCC 240
DB 217518 GCGAGTGCAGCAATAAAGAGTGCAGAGCTCAAGAAATAGCAGGTGCGGAAGCTAAGCC 217459

QY 241 TAAAGAACTTAAGACCGATTTCTGTAGAGGATGGAGCATCTTTCGTTCTGAGTGAATGC 300
DB 217458 TAAAGAACTTAAGACCGATTTCTGTAGAGGATGGAGCATCTTTCGTTCTGAGTGAATGC 217399

QY 301 TCTCATGCTCTGGCAGATAAGCTGGGTATTGTTCTAGTAACAGCTCGTCTTCTACTAG 360
DB 217398 TCTCATGCTCTGGCAGATAAGCTGGGTATTGTTCTAGTAACAGCTCGTCTTCTACTAG 217339

QY 361 CAGATCTGCAGACGTGGACTCAACGACGACGACCTACGCTCCCTCCACCCACGTT 420
DB 217338 CAGATCTGCAGACGTGGACTCAACGACGACGACCTACGCTCCCTCCACCCACGTT 217279

QY 421 TGATGATTAATAGACTCAAGCGCAACACAGCTTAACGATACTATCTTTACCTCAACATCACT 480
DB 217278 TGATGATTAATAGACTCAAGCGCAACACAGCTTAACGATACTATCTTTACCTCAACATCACT 217219

QY 481 AGCTGACATACAGGCTGCTTTCGTGAGCTCCAGGATGCTGCTCACTAATAATAAGATAC 540
DB 217218 AGCTGACATACAGGCTGCTTTCGTGAGCTCCAGGATGCTGCTCACTAATAATAAGATAC 217159

QY 541 AGCGGCTACTGATGAGGAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC 600
DB 217158 AGCGGCTACTGATGAGGAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC 217099

QY 601 AGTTAAAGTTGGCGGCAAAATACAGAAATAGCGAAATATGCTTTCGATTAACCAAGCGAT 660
DB 217098 AGTTAAAGTTGGCGGCAAAATACAGAAATAGCGAAATATGCTTTCGATTAACCAAGCGAT 217039

QY 661 TCTTGACTCTTTAGGTAAACTCACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 720
DB 217038 TCTTGACTCTTTAGGTAAACTCACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 216979

QY 721 TCTAGCAAACTAAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAAGTACT 780
DB 216978 TCTAGCAAACTAAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAAGTACT 216919

QY 781 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGAC 840
DB 216918 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGAC 216859

QY 841 ACAGATAGAGAAAGATGGAATGCGATTAGGATGATATTTTTCAGGACAGACGCTAG 900
DB 216858 ACAGATAGAGAAAGATGGAATGCGATTAGGATGATATTTTTCAGGACAGACGCTAG 216799

QY 901 TCGAGCTGTAGAAAATGCTTAAATCTTAAACAGATTAAGCAACATAGATTTCAGCTTAAAGC 960
DB 216798 TCGAGCTGTAGAAAATGCTTAAATCTTAAACAGATTAAGCAACATAGATTTCAGCTTAAAGC 216739

QY 961 AGCAATCGCTACTGTCTAAGACACAAATAGCTGAGCTCAGAAAAAGTTCCCGCATCTTCC 1020
DB 216738 AGCAATCGCTACTGTCTAAGACACAAATAGCTGAGCTCAGAAAAAGTTCCCGCATCTTCC 216679

QY 1021 AATTCTTCAAGAAAGCGGAACAAATCGTAAATACAGCTCAGAAAGATCTTAAAAAATATCA 1080
```


Db 216678 AATTTCTCAGGAAGCGGAAACAATGGTATATACAGGCTGAGAAGATCTTAAAAATATCAA 216619
QY 1081 ACCTGCAGATGGTTCTGATGTTTCCAAATCAGGAATACAGTTGGAGGCTCCAAAGCAACA 1140
Db 216618 ACCTGCAGATGGTTCTGATGTTTCCAAATCAGGAATACAGTTGGAGGCTCCAAAGCAACA 216559
QY 1141 AGGAAGTAGATATGGTAGTATTCGGTGTTCCTAGTCTTTAGATGATGCTGAAAATGAGAC 1200
Db 216558 AGGAAGTAGATATGGTAGTATTCGGTGTTCCTAGTCTTTAGATGATGCTGAAAATGAGAC 216499
QY 1201 CGCTTCCATTTTATGCTCGGGTTTCGTCAGATGATTCACATGTTCAATACGAAAAATCC 1260
Db 216498 CGCTTCCATTTTATGCTCGGGTTTCGTCAGATGATTCACATGTTCAATACGAAAAATCC 216439
QY 1261 TGATTCCTAAGCTGCCAACAGGAGCTCCGACGACAAGCTAGACAGCGAAAAGCGCTGG 1320
Db 216438 TGATTCCTAAGCTGCCAACAGGAGCTCCGACGACAAGCTAGACAGCGAAAAGCGCTGG 216379
QY 1321 AGATGACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGG 1380
Db 216378 AGATGACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGG 216319
QY 1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCT 1440
Db 216318 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCT 216259
QY 1441 TGTGAGCGGAGAGTTCTCCCGCTCGAGCAAGTTCTATAGGGTCAATCTGTAACAGCT 1500
Db 216258 TGTGAGCGGAGAGTTCTCCCGCTCGAGCAAGTTCTATAGGGTCAATCTGTAACAGCT 216199
QY 1501 TTACAGAGCTCCTCAAAATCTACAGGTTCTGATTAATAAAACACAGATATCAGAGGTTATCA 1560
Db 216198 TTACAGAGCTCCTCAAAATCTACAGGTTCTGATTAATAAAACACAGATATCAGAGGTTATCA 216139
QY 1561 TGCTTCAAAATCATCAATGATGCTCTATGTTAGGCGACGAAATGATCGACTCGTGAATG 1620
Db 216138 TGCTTCAAAATCATCAATGATGCTCTATGTTAGGCGACGAAATGATCGACTCGTGAATG 216079
QY 1621 GATAAACAATGTAAGTACCCCGCTCTCACAGATCGGTTCTTAGAGCAGCAGACGAAAGC 1680
Db 216078 GATAAACAATGTAAGTACCCCGCTCTCACAGATCGGTTCTTAGAGCAGCAGACGAAAGC 216019
QY 1681 TCAGAGCAGCAGAAAAACAGATCAAGCCCTCGTAGGTTGATTTCTGGCAATAGCAGAAAC 1740
Db 216018 TCAGAGCAGCAGAAAAACAGATCAAGCCCTCGTAGGTTGATTTCTGGCAATAGCAGAAAC 215959
QY 1741 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTTACAAATCTGTAATCGAGATCATCCAGTC 1800
Db 215958 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTTACAAATCTGTAATCGAGATCATCCAGTC 215899
QY 1801 GAATCTCAGCGCAATTAATGAGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCC 1860
Db 215898 GAATCTCAGCGCAATTAATGAGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCC 215839
QY 1861 TCACAGATTTGGTATCTCTATGTCAACTTTCTAATGACTCTTACAGAGTTCTATAGC 1920
Db 215838 TCACAGATTTGGTATCTCTATGTCAACTTTCTAATGACTCTTACAGAGTTCTATAGC 215779
QY 1921 TAAATTAGAAAGTTTGTGCTGAAGATCTAGGACAGGAGCTGAATTAAGACATTTTC 1980
Db 215778 TAAATTAGAAAGTTTGTGCTGAAGATCTAGGACAGGAGCTGAATTAAGACATTTTC 215719
QY 1981 CTTTGAACCAACTCTCTGTTTATTCAGCAGGTTCTGTTCAATATCGGCTCTCTATATTC 2040
Db 215718 CTTTGAACCAACTCTCTGTTTATTCAGCAGGTTCTGTTCAATATCGGCTCTCTATATTC 215659
QY 2041 TGGTTATCTCCAATAAACACACCTTAAGTGTTCGTTTGGAGAGATTAATATGCTTTGGT 2100
Db 215658 TGGTTATCTCCAATAAACACACCTTAAGTGTTCGTTTGGAGAGATTAATATGCTTTGGT 215599
QY 2101 AAGGCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATAAATAAAGA 2156

Db 215598 AAGGCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 215543
RESULT 6
AR310754_08/c
WPCOMMENT
Sequence split into 13 fragments LOCUS AR310754 Accession AR310754
Fragment Name Begin End
AR310754_00 1 110000
AR310754_01 100001 210000
AR310754_02 200001 310000
AR310754_03 300001 410000
AR310754_04 400001 510000
AR310754_05 500001 610000
AR310754_06 600001 710000
AR310754_07 700001 810000
AR310754_08 800001 910000
AR310754_09 900001 1010000
AR310754_10 1000001 1110000
AR310754_11 1100001 1210000
AR310754_12 1200001 1230025
Continuation (9 of 13) of AR310754 from base 800001 (AR310754 Sequence 1 from patent US
Query Match 99.4%; Score 2143.4; DB 6; Length 110000;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2155; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 AATAAATCTTTAAACACAGGCTCCGATTAATATATAGTGAGAGCTTTTATTTTATTTT 60
Db 28698 AATAAATCTTTAAACACAGGCTCCGATTAATATATAGTGAGAGCTTTTATTTTATTTT 28639
QY 61 AATAAATACTAAACAGATTTTATTTATTTTATGTTTTTATGTTTATGTTTAACTTATTTT 120
Db 28638 AATAAATACTAAACAGATTTTATTTATTTTATGTTTTTATGTTTATGTTTAACTTATTTT 28579
QY 121 AGGTCTTATAGCAAGAAACAGAACGACACCTCCCGCAGATCTTTCTGCTCAAGATTGA 180
Db 28578 AGGTCTTATAGCAAGAAACAGAACGACACCTCCCGCAGATCTTTCTGCTCAAGATTGA 28519
QY 181 GCGAGTGCAGCAATAAGAGTGGGAGAGCTCAAAAGATAGCAGGTGGGAGCTTAAGCC 240
Db 28518 GCGAGTGCAGCAATAAGAGTGGGAGAGCTCAAAAGATAGCAGGTGGGAGCTTAAGCC 28459
QY 241 TAAAGAACTTAAGACCGATTCTGTAGAGCGATGGAGCATCTTTCGTTCTGCAAGTAATGC 300
Db 28458 TAAAGAACTTAAGACCGATTCTGTAGAGCGATGGAGCATCTTTCGTTCTGCAAGTAATGC 28399
QY 301 TCTCATGATCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAAACAGCTCGTCTTCTACTAG 360
Db 28398 TCTCATGATCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAAACAGCTCGTCTTCTACTAG 28339
QY 361 CAGATCTGCAGACGTTGACTCAACACAGCGACCGCAGCTACGCTCTCCACCCACGTT 420
Db 28338 CAGATCTGCAGACGTTGACTCAACACAGCGACCGCAGCTACGCTCTCCACCCACGTT 28279
QY 421 TCATGATTTATAGACTCAAGCGCAACAGCTTACGATATATCTTTCCTCAACATCACT 480
Db 28278 TCATGATTTATAGACTCAAGCGCAACAGCTTACGATATATCTTTCCTCAACATCACT 28219
QY 481 AGCTGACATACAGGCTGCTTTGTCAGCTCCAGGATCTGTCACATAATAAAGGATAC 540
Db 28218 AGCTGACATACAGGCTGCTTTGTCAGCTCCAGGATCTGTCACATAATAAAGGATAC 28159
QY 541 ACGGCTCTACTGATAGGAAACCGCAATCGCTGCGAGTGGGAAACTTAAAGATTCGCGATGC 600
Db 28158 ACGGCTCTACTGATAGGAAACCGCAATCGCTGCGAGTGGGAAACTTAAAGATTCGCGATGC 28099
QY 601 AGTTAAAGTTGGCGCAAAATTAAGAAATATGCTTCGATTAACCAAGCGAT 660
Db 28098 AGTTAAAGTTGGCGCAAAATTAAGAAATATGCTTCGATTAACCAAGCGAT 28039
QY 661 TCTTGACTCTTTAGGTAAACTGACTTCTTCTGAGGCTCTTCTTCTCAATC 720

Db 28038 TCTTGACTCTTTAGTAAAGTGAAGTCTCTTTCGAGCTCTTACAGGCTGCTCTCTTCCAAATC 27979
Qy 721 TGTAGCAAAACATTAACAAGCAGCTGAGCTCTTTAAAGAGATGACAGATACCCAGTAGT 780
Db 27978 TGTAGCAAAACATTAACAAGCAGCTGAGCTCTTTAAAGAGATGACAGATACCCAGTAGT 27919
Qy 781 CCCAGGAAAACGGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGAC 840
Db 27918 CCCAGGAAAACGGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGAC 27959
Qy 841 ACAGATAGAAAAGATGGAATGCGATTTAGGATGCAATATTTTTCAGGACAGACAGCTAG 900
Db 27858 ACAGATAGAAAAGATGGAATGCGATTTAGGATGCAATATTTTTCAGGACAGACAGCTAG 27799
Qy 901 TGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATTAAGCAACATAGATTCAGCTAAAGC 960
Db 27798 TGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATTAAGCAACATAGATTCAGCTAAAGC 27739
Qy 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCC 1020
Db 27738 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCC 27679
Qy 1021 AATTCTTCAAGAGCGGAACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
Db 27678 AATTCTTCAAGAGCGGAACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 27619
Qy 1081 ACCTGCAGATGTTCTGATGTTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACA 1140
Db 27618 ACCTGCAGATGTTCTGATGTTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACA 27559
Qy 1141 AGGAAGTAGTATTGGTAGTATTTCGTTGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
Db 27558 AGGAAGTAGTATTGGTAGTATTTCGTTGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 27499
Qy 1201 CGTTTCCATTTTGTATGCTGCGTTTTCGTCAGATGATTCATGTTTCAATACGGAATATCC 1260
Db 27498 CGTTTCCATTTTGTATGCTGCGTTTTCGTCAGATGATTCATGTTTCAATACGGAATATCC 27439
Qy 1261 TGATTTCAAGCTGCCAACAGGAGCTCGCAGACACAAGCTAGCAGGAAAGCGCTGG 1320
Db 27438 TGATTTCAAGCTGCCAACAGGAGCTCGCAGACACAAGCTAGCAGGAAAGCGCTGG 27379
Qy 1321 AGATGACAGTGTCTGTCAGGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTTAGG 1380
Db 27378 AGATGACAGTGTCTGTCAGGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTTAGG 27319
Qy 1381 TAAAGCTGGGCAACACAGGGGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTG- 1439
Db 27318 TAAAGCTGGGCAACACAGGGGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGT 27259
Qy 1440 TTGTGAGCGCAGGAGTCTCTCCGCTGACAGAAAGTTCTATAGGGTCATCTGTAAAACAGC 1499
Db 27258 TTGTGAGCGCAGGAGTCTCTCCGCTGACAGAAAGTTCTATAGGGTCATCTGTAAAACAGC 27199
Qy 1500 TTTTCAAGACCTCAAAATCTACAGTTCTGATTTATAAACAACAGATATCAGCAGGTTATG 1559
Db 27198 TTTTCAAGACCTCAAAATCTACAGTTCTGATTTATAAACAACAGATATCAGCAGGTTATG 27139
Qy 1560 ATGCTTACAAATCCATCAATGATGCTATGTTAGGGGCGAATAATGATGCGACTCTGATG 1619
Db 27138 ATGCTTACAAATCCATCAATGATGCTATGTTAGGGGCGAATAATGATGCGACTCTGATG 27079
Qy 1620 TGATTAACAATCTAAGTACCCCGCTCTCACAGATTCGTTTCTTAGAGCAGCAACAGAG 1679
Db 27078 TGATTAACAATCTAAGTACCCCGCTCTCACAGATTCGTTTCTTAGAGCAGCAACAGAG 27019
Qy 1680 CTCGAGGACCAAGAAAACAGATCAAGGCTCGCTAGGGTATGTTTCTGGCAATAGCAGAA 1739
Db 27018 CTCGAGGACCAAGAAAACAGATCAAGGCTCGCTAGGGTATGTTTCTGGCAATAGCAGAA 26959
Qy 1740 CTCCTGGAGATGCTATAGTCAAGTTTTCGGCACTCAATCTGTATGAGATCATCCAGT 1799
Db 26958 CTCCTGGAGATGCTATAGTCAAGTTTTCGGCACTCAATCTGTATGAGATCATCCAGT 26899

Qy 1800 CGAATCTCAAGCGAATAATGAGAGATCAGACAAAAGCTTATCATCGGAGTGACAAGC 1859
Db 26898 CGAATCTCAAGCGAATAATGAGAGATCAGACAAAAGCTTATCATCGGAGTGACAAGC 26839
Qy 1860 CTCACAGTTTGGCTATCCTTATGTGCAACTTTCTAATGATCTTACACAGAAGTTCTATAG 1919
Db 26838 CTCACAGTTTGGCTATCCTTATGTGCAACTTTCTAATGATCTTACACAGAAGTTCTATAG 26779
Qy 1920 CTAAATTTAGAAAGTTTGTGCTGAAGGATCAGGACAGAGCTGAATATAAAGCACTTT 1979
Db 26778 CTAAATTTAGAAAGTTTGTGCTGAAGGATCAGGACAGAGCTGAATATAAAGCACTTT 26719
Qy 1980 CCTTTGAAAACGAACTCCTCTGTTTATTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATT 2039
Db 26718 CCTTTGAAAACGAACTCCTCTGTTTATTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATT 26659
Qy 2040 CTGTTTATCTCCAAATAACAACACTTAAGTGTTCGTTTGGAGAGATTTATTTGCTTTGG 2099
Db 26658 CTGTTTATCTCCAAATAACAACACTTAAGTGTTCGTTTGGAGAGATTTATTTGCTTTGG 26599
Qy 2100 TAAAGCCCTTTGTTGAGGCTTACCAACACACTAGACGATCTTCAATAATAAAGA 2156
Db 26598 TAAAGCCCTTTGTTGAGGCTTACCAACACACTAGACGATCTTCAATAATAAAGA 26542

RESULT 7
AX338293
LOCUS AX338293 1956 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 63 from Patent WO0181379.
ACCESSION AX338293
VERSION AX338293.1 GI:18128828
KEYWORDS
SOURCE Chlamydothila pneumoniae
ORGANISM Chlamydothila pneumoniae
REFERENCE 1
AUTHORS Bhatia, A., Probst, P. and Stromberg, E. J.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0181379-A 63 01-NOV-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. .1956
/organism="Chlamydothila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"

ORIGIN
Query Match 90.7%; Score 1956; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ATGCTTAATCCTATTGTTCCAGGTCCTATAGACAAAACAGACGACACCTCCCGCAGAT 160
Db 1 ATGCTTAATCCTATTGTTCCAGGTCCTATAGACAAAACAGACGACACCTCCCGCAGAT 60
Qy 161 CTTTCTGCTCAAGGATTTGGAGGCGAGTGCAGCAATTAAGAGTGCAGAAAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTTGGAGGCGAGTGCAGCAATTAAGAGTGCAGAAAGCTCAAGAATA 120
Qy 221 GCAGGTGGGAAGCTAAGCCCTAAAGAACTTAAGACCAATCTGATGAGAGCGATGAGGATC 280
Db 121 GCAGGTGGGAAGCTAAGCCCTAAAGAACTTAAGACCAATCTGATGAGAGCGATGAGGATC 180
Qy 281 TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGCAGATAGCTGGGTATGCTTCTAGT 340
Db 181 TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGCAGATAGCTGGGTATGCTTCTAGT 240
Qy 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGCTGAGCTCAACGACAGCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGCTGAGCTCAACGACAGCGACCT 300

```
QY 401 ACGCTCTCCACCCACGTTTGTATGATATATAGACTCAAGCGCAAAACAGCTTACGATCT 460
Db 301 ACGCTCTCCACCCACGTTTGTATGATATATAGACTCAAGCGCAAAACAGCTTACGATCT 360
QY 461 ATCTTTACTCAACATCACTAGCTGACATACAGCTGCTTTGGTGAGCCTCCAGATGCT 520
Db 361 ATCTTTACTCAACATCACTAGCTGACATACAGCTGCTTTGGTGAGCCTCCAGATGCT 420
QY 521 GTCCTAATAATAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 580
Db 421 GTCCTAATAATAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 480
QY 581 GAACTAAGATCCGATGCGATTAAGCTTGGCGGCAAAATACAGATTAAGCAATAT 640
Db 481 GAACTAAGATGCGGATGCGATTAAGCTTGGCGGCAAAATACAGATTAAGCAATAT 540
QY 641 GCTTCGGATACCAAGCGATCTTGACTCTTTAGTTAACTGACTTCTTCGACTCTTA 700
Db 541 GCTTCGGATACCAAGCGATCTTGACTCTTTAGTTAACTGACTTCTTCGACTCTTA 600
QY 701 CAGGCTGCTCTTCTCCAACTCTGTAGCAAAATACAAAGAGCTGAGCTTCTTAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAACTCTGTAGCAAAATACAAAGAGCTGAGCTTCTTAAGAG 660
QY 761 ATGCAAGATACCCAGTAGTCCCGAGGAAACGCTGCAATTTGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATACCCAGTAGTCCCGAGGAAACGCTGCAATTTGCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCATAT 880
Db 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCATAT 780
QY 881 TTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATCAAGTATAGC 940
Db 781 TTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATCAAGTATAGC 840
QY 941 ACATAGATTCAGTTAAAGCAATCGCTACTGCTTAAGACAAATAGCTGAAGCTCAG 1000
Db 841 ACATAGATTCAGTTAAAGCAATCGCTACTGCTTAAGACAAATAGCTGAAGCTCAG 900
QY 1001 AAAAAGTTCCTCCAGCTCTCCAACTCTTCAAGAGCGGAACAAATGGTAAATCAGGCTGAG 1060
Db 901 AAAAAGTTCCTCCAGCTCTCCAACTCTTCAAGAGCGGAACAAATGGTAAATCAGGCTGAG 960
QY 1061 AAAGATCTTTAAATAATCAAACTGCAAGTGTCTGATGTTCCAAATCCAGGAACTACA 1120
Db 961 AAAGATCTTTAAATAATCAAACTGCAAGTGTCTGATGTTCCAAATCCAGGAACTACA 1020
QY 1121 GTTGAGGCTTCCAGCAACAGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTTA 1180
Db 1021 GTTGAGGCTTCCAGCAACAGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTTA 1080
QY 1181 GATGATGCTGAAAATAGACCGCTTCCATTTTGTATGCTGCGTTTTCGTCAGATGATTCAC 1240
Db 1081 GATGATGCTGAAAATAGACCGCTTCCATTTTGTATGCTGCGTTTTCGTCAGATGATTCAC 1140
QY 1241 ATGTTCAATACGGAATCTGATTTCTCAAGTGTCCACAGAGGCTGCGACACAGCT 1300
Db 1141 ATGTTCAATACGGAATCTGATTTCTCAAGTGTCCACAGAGGCTGCGACACAGCT 1200
QY 1301 AGAGCAGCGAAAGCGCTGGAGATGACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAA 1360
Db 1201 AGAGCAGCGAAAGCGCTGGAGATGACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAA 1260
QY 1361 GCTTTAGAACGGCTCTAGTTAAAGCTGGGCAACACAGGGGATATCTAATGCTTTAGGA 1420
Db 1261 GCTTTAGAACGGCTCTAGTTAAAGCTGGGCAACACAGGGGATATCTAATGCTTTAGGA 1320
QY 1421 CAGATCGCTTCTGCTGCTGTGTGAGCGCAGAGTTCTTCCCGCTGCGACAAAGTCTATA 1480
Db 1321 CAGATCGCTTCTGCTGCTGTGTGAGCGCAGAGTTCTTCCCGCTGCGACAAAGTCTATA 1380
QY 1481 GGGTCATCTGTAAAAACAGCTTTTACAGACCTCAAAATCTACAGGTTCTGTTATATAAAACA 1540
```

```
Db 1381 GGGTCATCTGTAAAAACAGCTTTTACAGACCTCAAAATCTACAGTTCTGATTTAAAAACA 1440
QY 1541 CAGATATCAGCAGGTTATGATGTTTACAAATCCATCAATGATGCTATGTTAGGCGACGA 1600
Db 1441 CAGATATCAGCAGGTTATGATGTTTACAAATCCATCAATGATGCTATGTTAGGCGACGA 1500
QY 1601 AATGATGCGACTGCTGATGATTAACAATGTAAGTACCCCGCTCTCACACGATCCGTT 1660
Db 1501 AATGATGCGACTGCTGATGATTAACAATGTAAGTACCCCGCTCTCACACGATCCGTT 1560
QY 1661 CCTAGACACGAACAGAGCTCGAGGACCAAGAAAACAGATCAAGCCCTCGCTAGGCTG 1720
Db 1561 CCTAGACACGAACAGAGCTCGAGGACCAAGAAAACAGATCAAGCCCTCGCTAGGCTG 1620
QY 1721 ATTTCTGCAATAGCAAACTCTTGGAGATGTCTATAGTCAAGTTTCGGCACTCAATCT 1780
Db 1621 ATTTCTGCAATAGCAAACTCTTGGAGATGTCTATAGTCAAGTTTCGGCACTCAATCT 1680
QY 1781 GTAATGAGATCATCCAGTCCGATCTTCAAGCGCAATTAATGAGGATCAGACAAAAGCTT 1840
Db 1681 GTAATGAGATCATCCAGTCCGATCTTCAAGCGCAATTAATGAGGATCAGACAAAAGCTT 1740
QY 1841 ACATCGGCACTGACAAAGCCTCCACAGTTTGGGTATCTTATGTCAACTTCTTAATGAC 1900
Db 1741 ACATCGGCACTGACAAAGCCTCCACAGTTTGGGTATCTTATGTCAACTTCTTAATGAC 1800
QY 1901 TCTACACAGAGTTCTAGCTAAATTAAGAAAGTTTGTCTGCTGAAGATCTAGGACAGCA 1960
Db 1801 TCTACACAGAGTTCTAGCTAAATTAAGAAAGTTTGTCTGCTGAAGATCTAGGACAGCA 1860
QY 1961 GCTCAAAATAAAGCACTTCTTGTGAACCACTCTTGTATTTCAGCAGGTCCTGCTC 2020
Db 1861 GCTCAAAATAAAGCACTTCTTGTGAACCACTCTTGTATTTCAGCAGGTCCTGCTC 1920
QY 2021 AATATCGGCTCTCTATATCTGTTTCTGTTTCTCAATAA 2056
Db 1921 AATATCGGCTCTCTATATCTGTTTCTGTTTCTCAATAA 1956
```

RESULT 8

AX349491
LOCUS AX349491 1956 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 14 from Patent WO202606.
ACCESSION AX349491
VERSION AX349491.1 GI:18615352
KEYWORDS

SOURCE

Chlamydomophila pneumoniae
Chlamydomophila pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

REFERENCE

1
Ratti, G. and Grandi, G.
Immunisation against Chlamydia pneumoniae
Patent: WO 0202606-A 14 10-JAN-2002;
Chiron S.p.A. (It)

FEATURES

source
1. 1956
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"

ORIGIN

Query Match 90.7%; Score 1956; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGTTAATCTCTATTTGGTCCAGGTCCTATAGCAAGAAACAGACCACTCCCGCAGAT 160
Db 1 ATGTTAATCTCTATTTGGTCCAGGTCCTATAGCAAGAAACAGACCACTCCCGCAGAT 60
QY 151 CTTTCTGCTCAAGATTTGAGCGAGTCAGCAATTAAGTGGGAGCTCAAGATATA 220
Db 61 CTTTCTGCTCAAGATTTGAGCGAGTCAGCAATTAAGTGGGAGCTCAAGATATA 120

221 GCAGGTGCGGAAGCTTAAGCTTAAGAAATCTAAGACCGATCTTGTAGAGCGATGGAGCATC 280
Db GCAGGTGCGGAAGCTTAAGCTTAAGAAATCTAAGACCGATCTTGTAGAGCGATGGAGCATC 180
281 TTGCGTTCTGCGAGTAATGCTCTCATGAGTCTGGCAGATAAGCTGGTATGCTTCTAGT 340
Db TTGCGTTCTGCGAGTAATGCTCTCATGAGTCTGGCAGATAAGCTGGTATGCTTCTAGT 240
341 AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGGGACTCAACGACAGCGACCGCACCT 400
Db AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGGGACTCAACGACAGCGACCGCACCT 300
401 ACSCCTCTCCACCCAGTTTGAATGATTAAGACTCAAGCGGCAACACAGCTTACGATCT 460
Db ACSCCTCTCCACCCAGTTTGAATGATTAAGACTCAAGCGGCAACACAGCTTACGATCT 360
461 ATCTTTACCTCAACATCAGCTAGCTGACATACAGGCTGCTTTGTGAGCCTCCAGGATGCT 520
Db ATCTTTACCTCAACATCAGCTAGCTGACATACAGGCTGCTTTGTGAGCCTCCAGGATGCT 420
521 GTCACCTAATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 580
Db GTCACCTAATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 480
581 GAAACTAAGATCCGATGCGATTAAAGTTGGCGCGCAAAATACAGAAATAGCGGAATAT 640
Db GAAACTAAGATCCGATGCGATTAAAGTTGGCGCGCAAAATACAGAAATAGCGGAATAT 540
641 GCTTCGGATAACCAAGGATCTTTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 700
Db GCTTCGGATAACCAAGGATCTTTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 600
701 CAGGCTGCTTCTTCCAACTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG 760
Db CAGGCTGCTTCTTCCAACTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG 660
761 ATCCAGATTAACCCAGTAGTCCAGGAAACCGCTGCAATCTCTCAATCTTTAGTGTAT 820
Db ATCCAGATTAACCCAGTAGTCCAGGAAACCGCTGCAATCTCTCAATCTTTAGTGTAT 720
821 CAGACAGATCTTACAGCGACACAGATAGAGAAAGATGGAATGCGAATAGGATGCAAT 880
Db CAGACAGATCTTACAGCGACACAGATAGAGAAAGATGGAATGCGAATAGGATGCAAT 780
881 TTTCAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGC 940
Db TTTCAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGC 840
941 AACATAGATTACGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
Db AACATAGATTACGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900
1001 AAAAAGTTCCCGACTCTCCAATCTTCAAGAGCGGAACAAATGGTAATACAGGCTGAG 1060
Db AAAAAGTTCCCGACTCTCCAATCTTCAAGAGCGGAACAAATGGTAATACAGGCTGAG 960
1061 AAGATCTTAAATAATCAACCTGCGAGATGGTTCTGATGTTCCAAATCAGGAACTPACA 1120
Db AAGATCTTAAATAATCAACCTGCGAGATGGTTCTGATGTTCCAAATCAGGAACTPACA 1020
1121 GTTCGAGCTCCAGCAACAGGAGTAGTATGCTAGTATTCGTGTTTCCATGCTCTTA 1180
Db GTTCGAGCTCCAGCAACAGGAGTAGTATGCTAGTATTCGTGTTTCCATGCTCTTA 1080
1181 GATGATGCTGAAAATGAGCCGCTTCCATTTTGTATGCTCGGTTTCGTAGATGATTCAC 1240
Db GATGATGCTGAAAATGAGCCGCTTCCATTTTGTATGCTCGGTTTCGTAGATGATTCAC 1140
1241 ATGTTCAATACGGAATTCCTGATCTCAAGCTGCCCAACAGGAGCTCGGACACAAGCT 1300
Db ATGTTCAATACGGAATTCCTGATCTCAAGCTGCCCAACAGGAGCTCGGACACAAGCT 1200

1301 AGACAGCGAAAGCCGCTGGAGATGACAGTGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1360
Db AGACAGCGAAAGCCGCTGGAGATGACAGTGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1260
1361 GCTTTAGAAAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGA 1420
Db GCTTTAGAAAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGA 1320
1421 CAGATCGCTTCTGCTGCTGTTGTAGCGCAGGNGTTCTCCGCTGCGCAGCAAGTCTATA 1480
Db CAGATCGCTTCTGCTGCTGTTGTAGCGCAGGAGTTCCTCCGCTGCGCAGCAAGTCTATA 1380
1481 GGGTCATCTGTAAAAACAGCTTTACAAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1540
Db GGGTCATCTGTAAAAACAGCTTTACAAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1440
1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGGTAGGCAAGA 1600
Db CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGGTAGGCAAGA 1500
1601 AATGATGCGACTCTGATGATGATTAACATGTAAGTACCCGCTCTCACAGATCCGTT 1660
Db AATGATGCGACTCTGATGATGATTAACATGTAAGTACCCGCTCTCACAGATCCGTT 1560
1661 CCTAGACAGCAACAGAGCTCGAGGACCAAGAAACAGATAAAGCAAGCCCTCGCTAGGGTG 1720
Db CCTAGACAGCAACAGAGCTCGAGGACCAAGAAACAGATAAAGCAAGCCCTCGCTAGGGTG 1620
1721 ATTCTGGCAATPAGCAGAACTCTTGAGATGCTATAGTCAAGTTTGGGCACTACAACT 1780
Db ATTCTGGCAATPAGCAGAACTCTTGAGATGCTATAGTCAAGTTTGGGCACTACAACT 1680
1781 GTAATGCAATCATCAGTCGAATCTCAAGCGAATATGAGGAGATCAGACAAAGCTT 1840
Db GTAATGCAATCATCAGTCGAATCTCAAGCGAATATGAGGAGATCAGACAAAGCTT 1740
1841 ACATCGGCACTGACAAAGCTCCAGATTGGCTATCTTATGTGCAATTTCTTAATGAC 1900
Db ACATCGGCACTGACAAAGCTCCAGATTGGCTATCTTATGTGCAATTTCTTAATGAC 1800
1901 TCTACACAGAGTTCTATAGCTAAATPAGAAAGTTTGTGCTGAGGATCTAGGACAGA 1960
Db TCTACACAGAGTTCTATAGCTAAATPAGAAAGTTTGTGCTGAGGATCTAGGACAGA 1860
1961 GCTGAAATAAAGCACTTCTTGAAGCAACTCTGTTTATTCAGAGGTGCTGCTC 2020
Db GCTGAAATAAAGCACTTCTTGAAGCAACTCTGTTTATTCAGAGGTGCTGCTC 1920
2021 AATATCGGCTCTCTATATCTGTTTATCTCCAATAA 2056
Db AATATCGGCTCTCTATATCTGTTTATCTCCAATAA 1956

RESULT 9
AX361992LOCUS AX361992 1956 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 385 from Patent WO0208267.

ACCESSION

AX361992

VERSION

AX361992.1

KEYWORDS

GI:18694447

SOURCE

Chlamydia pneumoniae

ORGANISM

Chlamydia pneumoniae

REFERENCE

1

AUTHORS

Fling, S.P., Skeiky, Y.A., Probst, P. and Bhatia, A.

TITLE

Compounds and methods for treatment and diagnosis of chlamydial infection

JOURNAL

Patent: WO 0208267-A 385 31-JAN-2002;

FEATURES

CORIXA CORPORATION (US)

Location/Qualifiers

1. .1956

/organism="Chlamydia pneumoniae"

/mol_type="unassigned DNA"

ORIGIN									
/db_xref="taxon:83558"									
Query Match 90.7%; Score 1956; DB 6; Length 1956;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	101	ATGGTAATCTCTATTGGTCCAGGTCTCTATAGACGAAACAGAACGACACCTCCCGCAGAT	160						
DB	1	ATGGTTAATCTCTATTGGTCCAGGTCTCTATAGACGAAACAGAACGACACCTCCCGCAGAT	60						
QY	161	CTTTCTGCTCAGAGATTGGAGCGGAGTGCAGACATTAAGATGAGGAGCGGAGCTCAAGAGATA	220						
DB	61	CTTTCTGCTCAAGGATTGGAGCGGAGTGCAGACATTAAGATGAGGAGCGGAGCTCAAGAGATA	120						
QY	221	GCAGGTGCGGAAGCTTAAGCTTAAGAACTTAAGACCGGATTCTGTAGAGCGATGAGAGCATC	280						
DB	121	GCAGGTGCGGAAGCTTAAGCTTAAGAACTTAAGACCGGATTCTGTAGAGCGATGAGAGCATC	180						
QY	281	TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGT	340						
DB	181	TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGT	240						
QY	341	AACAGCTGCTCTTACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCGACCT	400						
DB	241	AACAGCTGCTCTTACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCGACCT	300						
QY	401	ACGCTCTCTCCACCGTTTGATGATTATAAGATCAAGCGCAACAGCTTACGATATCT	460						
DB	301	ACGCTCTCTCCACCGTTTGATGATTATAAGATCAAGCGCAACAGCTTACGATATCT	360						
QY	461	ATCTTTACTCTACATCACTAGTGCACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT	520						
DB	361	ATCTTTACTCTACATCACTAGTGCACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT	420						
QY	521	GTCACTAATATAAGGATACAGCGGTACTGATGAGGAAACCGCAATCGCTCGGAGTGG	580						
DB	421	GTCACTAATATAAGGATACAGCGGTACTGATGAGGAAACCGCAATCGCTCGGAGTGG	480						
QY	581	GAAACTAAGATGCGATGCGATTAAAGTTGGCGGCAAAATACAGAAATAGCGAAATAT	640						
DB	481	GAAACTAAGATGCGATGCGATTAAAGTTGGCGGCAAAATACAGAAATAGCGAAATAT	540						
QY	641	GCTTGGGATAACCAAGCGATTCTTGACTCTTTAGGTAAAGTCACTTCTTCGACCTCTTA	700						
DB	541	GCTTGGGATAACCAAGCGATTCTTGACTCTTTAGGTAAAGTCACTTCTTCGACCTCTTA	600						
QY	701	CAGGCTGCTCTTCTCCAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG	760						
DB	601	CAGGCTGCTCTTCTCCAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG	660						
QY	761	ATGCAAGATAACCGAGTAGTCCAGGGAACAGCTGCAATTGCTCAATCTTTAGTTGAT	820						
DB	661	ATGCAAGATAACCGAGTAGTCCAGGGAACAGCTGCAATTGCTCAATCTTTAGTTGAT	720						
QY	821	CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGAATTAGGGATGCATAT	880						
DB	721	CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGAATTAGGGATGCATAT	780						
QY	881	TTTGGCAGGACAGAACCGCTAGTGGAGTGTAGAAAATGCTTAAATCTTAATAACAGATTAAGC	940						
DB	781	TTTGGCAGGACAGAACCGCTAGTGGAGTGTAGAAAATGCTTAAATCTTAATAACAGATTAAGC	840						
QY	941	AACATAGATTACGCTAAGAGCAGCAATCGCTACTGCTTAAGACACAATAGCTGAAGCTCAG	1000						
DB	841	AACATAGATTACGCTAAGAGCAGCAATCGCTACTGCTTAAGACACAATAGCTGAAGCTCAG	900						
QY	1001	AAAAGTTCCCGACTCTCCAATCTTCAAGAGCGGAACAATAGTAAATACAGGCTGAG	1060						
DB	901	AAAAGTTCCCGACTCTCCAATCTTCAAGAGCGGAACAATAGTAAATACAGGCTGAG	960						
QY	1061	AAAGATCTTAAATAATATCAAACTCGAGATGTTCTGATGTTCCAAATCCAGGAACCTACA	1120						

RESULT 10
AX045133
LOCUS AX045133
DEFINITION Sequence 3 from Patent WO0066739.
ACCESSION AX045133
VERSION AX045133.1 GI:11343732

1852 bp DNA linear PAT 24-NOV-2000


```
QY 1925 TTAGAAAGTTGTTCTGCTGAAGGATCTAGACAGCAGCTGAAATAAAGACATTTCTTTT 1984
Db 1621 TTAGAAGTTGTTGCTGAAGGATCTAGACAGCAGCTGAAATAAAGACATTTCTTTT 1680
QY 1985 GAAACGAATCCCTGTTTATTCAGCAGGCTGCTGGTCAATATCGGCTCTCTATATTCGTT 2044
Db 1681 GAAACGAATCCCTGTTTATTCAGCAGGCTGCTGGTCAATATCGGCTCTCTATATTCGTT 1740
QY 2045 TATCTCAATAAACAACACCTAAGTTCGTTGGAGAGATTATATGCTTTGGTAAGG 2104
Db 1741 TATCTCAATAAACAACACCTAAGTTCGTTGGAGAGATTATATGCTTTGGTAAGG 1800
QY 2105 CCTTTGTTAGGCTTACCAACACCTAGAACGATCTTCAATAAATAAAGA 2156
Db 1801 CCTTTGTTAGGCTTACCAACACCTAGAACGATCTTCAATAAATAAAGA 1852

RESULT 11
AX045137 2238 bp DNA linear PAT 24-NOV-2000
LOCUS
DEFINITION Sequence 7 from Patent WO0066739.
ACCESSION AX045137
VERSION AX045137.1 GI:11343736
KEYWORDS
SOURCE
ORGANISM
Chlamydomophila pneumoniae
Chlamydomophila pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
1
REFERENCE
Murdin, A.D., Omen, R.P., Wang, J. and Dunn, P.
AUTHORS i (Chlamydia) antigens and corresponding dna fragments and uses
TITLE thereof
JOURNAL Patent: WO 0066739-A 7 09-NOV-2000;
Aventis Pasteur Limited (CA)
FEATURES
source
location/Qualifiers
1. 2238
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
766..2238
/notes="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC17330.1"
/db_xref="GI:11343737"
/db_xref="REMTREMBL_CAC17330"
/translation="MNPPIGGPIDEPTPPADLSNQGLEASANKSABAQRIAGAE
AKPKESKIDSVERMSILURSVAVALMSLADKLGIASNSSSTSRSDVDSTTATPTP
PPFTDYKTAQAYDTIFTSLSLADIQALVSDAVTNIRKDTAATDETAIAEEL
ETKNADAVKVGAGITELAKYASDNQAILDSLGKLSFDLIQAALLOSVAANNKAAEL
KEMQDNVPVPGKTPALAQSLVDOTDATQIEKDGNARIDAYFAGQVAGSAGVENAKSN
NSTSIDSAXAATATQIAEAKKPPDPSPILQEBOMVIOAEKDKLNKPKADGSDV
PNPGTTVGSGKQSGSSIRVSMLLDDENETASILMSGPROXIHMENTENPDSQAA
QVELAQAARAKAGDSDSAALADAKALEALGKAGCQOQGIINALGQIASAAVSA
GVLPQLQVLRWIRQAIVSEQKLISEEDLNSAVDHHHHH"

Query Match 68.0%; Score 1465.8; DB 6; Length 2238;
Best Local Similarity 99.8%; Pred. No. 1.1e-310;
Matches 1478; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATAAAACTTTAAAAACAGGCTGCATTAAATATTAGTGAGAGCTTTTTTTTATTTTTT 60
Db 666 ATAAAACTTTAAAAACAGGCTGCATTAAATATTAGTGAGAGCTTTTTTTTATTTTTT 725
QY 61 ATATAAACTAAAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db 726 ATATAAACTAAAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 785
QY 121 AGGTCTTATAGACGAAACAGAACCCACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 180
Db 786 AGGTCTTATAGACGAAACAGAACCCACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 845
```

```
QY 181 GCGAGTGCAGCAAAATAAGAGTGCAGAAAGCTCAAAAGATAGCAGGTGCGGAAGCTAAGCC 240
Db 846 GCGAGTGCAGCAAAATAAGAGTGCAGAAAGCTCAAAAGATAGCAGGTGCGGAAGCTAAGCC 905
QY 241 TAAAGAACTAAGACCGGATTTCTGTAGAGCGATGGAGCATCTTGGCTTCGCGAGTGAATGC 300
Db 906 TAAAGAACTAAGACCGGATTTCTGTAGAGCGATGGAGCATCTTGGCTTCGCGAGTGAATGC 965
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAAACAGCTGCTTCTACTAG 360
Db 966 TCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAAACAGCTGCTTCTACTAG 1025
QY 361 CAGATCTGCAGACGCTGGACTCAACGACAGCGACCGCCTACGCTCTCCACCCACGTT 420
Db 1026 CAGATCTGCAGACGCTGGACTCAACGACAGCGACCGCCTACGCTCTCCACCCACGTT 1085
QY 421 TGAATGATTAAAGACTCAAGCGCAAAACAGCTTACGATATCTTTTCTCAACATCACT 480
Db 1086 TGAATGATTAAAGACTCAAGCGCAAAACAGCTTACGATATCTTTTCTCAACATCACT 1145
QY 481 AGCTGACATACAGGCTGCTTTGCTGAGCCTCCAGGATGCTGCTCACTAAATAAAGATAC 540
Db 1146 AGCTGACATACAGGCTGCTTTGCTGAGCCTCCAGGATGCTGCTCACTAAATAAAGATAC 1205
QY 541 AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC 600
Db 1206 AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC 1265
QY 601 AGTTAAAGTTGGCGGCAAAATACAGAAATAGGAAATATGCTTCGGATAACCAAGCGAT 660
Db 1266 AGTTAAAGTTGGCGGCAAAATACAGAAATAGGAAATATGCTTCGGATAACCAAGCGAT 1325
QY 661 TCTTGAATCTTTAGGTAAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 720
Db 1326 TCTTGAATCTTTAGGTAAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 1385
QY 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db 1386 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 1445
QY 781 CCCAGGAAACCCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db 1446 CCCAGGAAACCCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 1505
QY 841 ACAGATAGAGAAAGATGGAAATGCGGATTTAGGATGCAATTTTGCAGGACAGAACCCCTAG 900
Db 1506 ACAGATAGAGAAAGATGGAAATGCGGATTTAGGATGCAATTTTGCAGGACAGAACCCCTAG 1565
QY 901 TGGAGCTGTAGAAAATGCTTAATATCTAATAACAGTATAAGCAACATAGATTAGCTAAAGC 960
Db 1566 TGGAGCTGTAGAAAATGCTTAATATCTAATAACAGTATAAGCAACATAGATTAGCTAAAGC 1625
QY 961 AGCAATCGCTACTGCTTAAGCAACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
Db 1626 AGCAATCGCTACTGCTTAAGCAACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1685
QY 1021 AATTCTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATATCAA 1080
Db 1686 AATTCTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATATCAA 1745
QY 1081 ACTGCGAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACA 1140
Db 1746 ACTGCGAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACA 1805
QY 1141 AGAAGTGTATTTAGTATTTGCTGTTTCCATGCTGTTAGTATGATGCTGAAATCAGAC 1200
Db 1806 AGAAGTGTATTTAGTATTTGCTGTTTCCATGCTGTTAGTATGATGCTGAAATCAGAC 1865
QY 1201 CGCTTCCATTTTGAATGCTCTGGGTTTCCGTCAGATGATTCATGTTCAATACGGAATATCC 1260
Db 1866 CGCTTCCATTTTGAATGCTCTGGGTTTCCGTCAGATGATTCATGTTCAATACGGAATATCC 1925
QY 1261 TGAATCTCAAGTGCACCAACAGGAGCTCGCAGCACCAAGCTAGAGCGGAAACCCGCTGG 1320
```

Db 1926 TGATTTCTTAAGTGTCCCAACGAGGAGTCGCAGCACAAAGCTTAGACGACGAAGCGCGTG 1985

QY 1321 AGATGACAGTGTCTGCTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGCTCTPAGG 1380

Db 1986 AGATGACAGTGTCTGCTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGCTCTPAGG 2045

QY 1381 TAAAGCTGGGCAACAACAGGCCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440

Db 2046 TAAAGCTGGGCAACAACAGGCCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 2105

QY 1441 TGTGAGCGCAGGAGTTCCTCCCGCTGCAGCAAGTTCCTATAG 1481

Db 2106 TGTGAGCGCAGGAG-TACTCCCCTGCAGCAAGTTCCTATG 2145

RESULT 12
CHT76KDA

LOCUS CHT76KDA 2148 bp DNA linear BCT 17-MAR-1994
DEFINITION Chlamydia pneumoniae 76 kDa protein gene, complete cds.
ACCESSION L23921
VERSION L23921.1 GI:435961
KEYWORDS 76 kDa protein;
SOURCE Chlamydophila pneumoniae
ORGANISM Chlamydophila pneumoniae
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales; Chlamydiales; Chlamydiales
REFERENCE 1 (bases 1 to 2148);
AUTHORS Perez Melgosa, M., Kuo, C.C. and Campbell, L.A.
TITLE Isolation and characterization of a gene encoding a Chlamydia
pneumoniae 76-kilodalton protein containing a species-specific
epitope
JOURNAL Infect. Immun. 62 (3), 880-886 (1994)
MEDLINE 94156481
PUBMED 7509320
COMMENT Original source text: Chlamydia pneumoniae DNA.
FEATURES
source Location/Qualifiers
1..2148 /organism="Chlamydophila pneumoniae"
/mol_type="genomic DNA"
/db_xref="taxon:83558"
1..2148 /codon_start=1
/transl_table=11
/product="76 kDa protein"
/protein_id="A223117.1"
/db_xref="GI:435962"
/translation="MTKKHVVWVEGLNRLPKQFVKGSVVDNVTVPSETSTTEKA
ATMAKYVCVMVLGVKSQVPIWGKPLLYLGFALMNPFLAKATKTLLNGKENLA
WFIGTLGLRKGADMSAVRYEVALSVPEIDVGIRGNLLKEFWFAQAATAANDYP
KSNFGNTNYKFGLAWYGITDLSFRAYGASKPANDKLGSDFTFKFDLGLISAFG
KNFNKIIFKNRLAIIESFFFIIFYNKTFELLFFELVAPNIPGPDIETERTPDPA
LSAQGLEASANAABQAQTAGAEAKPKESKTVERWSLTRSAVNALLSLADKLUGIA
SSNSGSTSRSDASTATAPPPTPPDDYKTAQTDVIETSTSLADIQAALVS
LDADVNIKDTHATDEBTALAWEYNKADAVKVGAQIHELAKYADNQAILDSLGKL
TSFDILPAQLQSVAANNNAEKLEKMOPNVPGKTPALAOGLSDVDQDATATPQIEK
GNRIIDAYFAGQQSAVENNAENNNISINDSKAAIAATAKTIQIAEAOKFPDPSPILQ
EARQVITYAEKDLIKTKPADGPSVNPVPTTVGGSGKOOGSSI GSI RYSMLDLLDAENETA
SIUMSGFMWHMFNTENPDSOAQCQLAAQARAACAAGDSDSAAAAAADQAQALEAAL
GKAGQQQGIIINALGQIASAAVVSAGVIPQOVL"

CDS

Query Match 67.6%; Score 1458; DB 1; Length 2148;
Best Local Similarity 99.9%; Pred. No. 5.9e-309;
Matches 1480; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 ATAAATCTTTAAAAACAGGCTCGCATTAATTATAGTAGAGAGCTTTTTTTATTTTTTT 60

Db 668 ATAATATCTTTAAAAAACAGGCTCGCATTAATTATAGTAGAGAGCTTTTTTTATTTTTTT 727

QY 61 ATAATAAACCTAAAGAGATTTTTTATTTTTTCAGTTTTTA-TGGTTATCCCTATTGGTC 119

Db 728 ATAATAAACCTAAAGAGATTTTTTATTTTTTCAGTTTTTTTATTTGGTTAAFCCTATTGGTC 787


```
Db      1868  CCGCTTCATTTGATCTGGGTTTCGTGAGATGATTCATGTTCAATACGGAATC 1927
Qy      1260  CTGATTTCTAAGCTGCCAACAGAGAGCTCGCAGCACAAAGCTAGAGCAGCGAAAGCCGCTG 1319
Db      1928  CTGATTTCTAAGCTGCCAACAGAGAGCTCGCAGCACAAAGCTAGAGCAGCGAAAGCCGCTG 1987
Qy      1320  GAGATGACAGTCTGCTGCGAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAG 1379
Db      1988  GAGATGACAGTCTGCTGCGAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAG 2047
Qy      1380  GTAAGCTGGGCAACACAGGGCATCTCAATGCTTTAGACAGATCGCTTCTGCTCTG 1439
Db      2048  GTAAGCTGGGCAACACAGGGCATCTCAATGCTTTAGACAGATCGCTTCTGCTCTG 2107
Qy      1440  TTGTGAGCGGAGGATTCCTCCGCTGCAGCAAGTTCTATAG 1481
Db      2108  TTGTGAGCGGAGGAG-TCTCCGCTGCAGCAAGTTCTATAG 2148

RESULT 13
AX045135
LOCUS      1456 bp      DNA      linear      PAT 24-NOV-2000
DEFINITION      Sequence 5 from Patent WO0066739.
ACCESSION      AX045135
VERSION      AX045135.1 GI:11343734
KEYWORDS
SOURCE      Chlamydomonadales
ORGANISM      Chlamydomonadales; Chlamydiales; Chlamydiaceae; Chlamydomonadales.
REFERENCE      1
AUTHORS      Murdin, A.D., Omen, R.P., Wang, J. and Dunn, P.
TITLE      i(chlamydia) antigens and corresponding dna fragments and uses
JOURNAL      Patent: WO 0066739-A 5 09-NOV-2000;
FEATURES
source      1..1456
            /organism="Chlamydomonadales"
            /mol_type="unassigned DNA"
            /db_xref="taxon:83558"
            101..>1452
            /note="unnamed protein product"
            /codon_start=1
            /transl_table=11
            /protein_id="CAC17379.1"
            /db_xref="GI:11343735"
            /db_xref="REMTREMBL:CAC17379"
            /translation="MVPNIGPPIDETPPPADLSAQGLEASAAKSAEQAQRIAGAE
            AKPKESKTSVERWSILRSVNAALMSLADKLGIASSSSSTSRSDVDSTTATPTP
            PPTFDYKTOAGTAYDTITFTSLADIQALVSLQDAVNIKDTATDETAIAEW
            ETKNADAVKQAGITELAKVASDNQALLSLGKLTSPDLQALLQSVANNKAEEL
            KEMQNPVPEKPAIAQSLVDQDATQIEKGNADIRDAYFAGNAGAVENAKSN
            NSISNDSAKAAIAKTQIAEAKKFPDPSPILOEAPQWVIOAKDKLKNIKPADGSDV
            PNFTVTGGSKQSSSTSRVSLMLDAENETASILMSFGMIFMNFENPDSDAA
            QOELAAQARAAGAAGDSDAAALADAKALEALGKAGQQQGLNALGQIASAAVVA
            G"

ORIGIN
Query Match      67.5%; Score 1455; DB 6; Length 1456;
Best Local Similarity 100.0%; Pred. No. 2.8e-308;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  ATAAAATCTTTAAAACAGGCTCCGATTAATATTAGTGAGAGGCTTTTATTTT 60
Db      1  ATAAAATCTTTAAAACAGGCTCCGATTAATATTAGTGAGAGGCTTTTATTTT 60
Qy      61  ATATAAACTAAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db      61  ATATAAACTAAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Qy      121  AGTCTCTATAGACGAAACAGAACGCAACCTCCGAGATCTTCTGCTCAAGGATGGA 180
```

```
Db      121  AGTCTCTATAGACGAAACAGAACGCAACCTCCGAGATCTTCTGCTCAAGGATGGA 180
Qy      181  GCGAGCTGACGAAATTAAGAGTCCGGAAGCTCAAGAATAGCAGGTGCGGAAGTAAAGCC 240
Db      181  GCGAGCTGACGAAATTAAGAGTCCGGAAGCTCAAGAATAGCAGGTGCGGAAGTAAAGCC 240
Qy      241  TAAAGAACTTAAGACCGATTTCTGTAGAGGATGGAACATCTTGCCTTCTGAGTGAATGC 300
Db      241  TAAAGAACTTAAGACCGATTTCTGTAGAGGATGGAACATCTTGCCTTCTGAGTGAATGC 300
Qy      301  TCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAAACAGCTCGTCTTCTACTAG 360
Db      301  TCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAAACAGCTCGTCTTCTACTAG 360
Qy      361  CAGATCTGACAGCTGGAATCTCAACGACAGGACCGCACTACGCTTCTCCACCCAGCTT 420
Db      361  CAGATCTGACAGCTGGAATCTCAACGACAGGACCGCACTACGCTTCTCCACCCAGCTT 420
Qy      421  TGAATGATTATAAGACTCAAGCGCAACAGCTTACGATCTATCTTTTACCTCAACATCACT 480
Db      421  TGAATGATTATAAGACTCAAGCGCAACAGCTTACGATCTATCTTTTACCTCAACATCACT 480
Qy      481  AGCTGACATACAGGCTGCTTTGCTGAGCTTCCAGGATGCTGCTCACTAATAAAGATAC 540
Db      481  AGCTGACATACAGGCTGCTTTGCTGAGCTTCCAGGATGCTGCTCACTAATAAAGATAC 540
Qy      541  AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAACTAAGAAATGCCGATGC 600
Db      541  AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAACTAAGAAATGCCGATGC 600
Qy      601  AGTAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT 660
Db      601  AGTAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT 660
Qy      661  TCTTGAATCTTTAGGTAATGCTGCTTCTTGAAGCTTCTTACAGGCTGCTTCTTCTCAATC 720
Db      661  TCTTGAATCTTTAGGTAATGCTGCTTCTTGAAGCTTCTTACAGGCTGCTTCTTCTCAATC 720
Qy      721  TGTAGCAAAATACAAAGAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTACT 780
Db      721  TGTAGCAAAATACAAAGAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTACT 780
Qy      781  CCCAGGAAAAAGCCCTGCAATGCTCAATCTTTAGTTGATGACAGAGATGCTACAGGAC 840
Db      781  CCCAGGAAAAAGCCCTGCAATGCTCAATCTTTAGTTGATGACAGAGATGCTACAGGAC 840
Qy      841  ACAGATAGAGAAAGATGGAATGCGAATGAGGATGCAATTTTTCAGAGCAGAACGCTAG 900
Db      841  ACAGATAGAGAAAGATGGAATGCGAATGAGGATGCAATTTTTCAGAGCAGAACGCTAG 900
Qy      901  TGGAGCTGTAGAAATGCTTAAATCTAATACAGTATAAGCAACATAGATTCAGTAAAGC 960
Db      901  TGGAGCTGTAGAAATGCTTAAATCTAATACAGTATAAGCAACATAGATTCAGTAAAGC 960
Qy      961  AGCAATCGCTACTGCTTAAGACACAAATAGCTGAGCTCAGAAAAAGTTCCCGACTCTCC 1020
Db      961  AGCAATCGCTACTGCTTAAGACACAAATAGCTGAGCTCAGAAAAAGTTCCCGACTCTCC 1020
Qy      1021  AATCTTCAAGAAAGCGGAACAAATGTAATACNGGCTGAGAAAGATCTTAAAAATATCAA 1080
Db      1021  AATCTTCAAGAAAGCGGAACAAATGTAATACNGGCTGAGAAAGATCTTAAAAATATCAA 1080
Qy      1081  ACCTGAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 1140
Db      1081  ACCTGAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 1140
Qy      1141  AGGAAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAATGAGAC 1200
Db      1141  AGGAAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAATGAGAC 1200
Qy      1201  CGCTTCCATTTTGAATGCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATTC 1260
Db      1201  CGCTTCCATTTTGAATGCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATTC 1260
```

QY 1261 TGATTCTCAAGCTGCCAACAGGAGCTCGACAGCAAGCTAGACGACGAAAGCCGCTGG 1320
 Db 1261 TGATTCTCAAGCTGCCAACAGGAGCTCGACAGCAAGCTAGACGACGAAAGCCGCTGG 1320
 QY 1321 AGATGACAGTCTGCTCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
 Db 1321 AGATGACAGTCTGCTCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
 QY 1381 TAAAGCTGGGCAACAAAGGCGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
 Db 1381 TAAAGCTGGGCAACAAAGGCGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
 QY 1441 TGTGAGCGCAGGAGT 1455
 Db 1441 TGTGAGCGCAGGAGT 1455

RESULT 14
 LOCUS AE016994 304769 bp DNA linear BCT 15-AUG-2003
 DEFINITION Chlamydomophila caviae GPIC section 1 of 4 of the complete genome.
 ACCESSION AE016994 AE015925
 VERSION AE016994.1 GI:29834116
 KEYWORDS
 SOURCE Chlamydomophila caviae GPIC
 ORGANISM Chlamydomophila caviae GPIC
 Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomphila.
 REFERENCE 1 (bases 1 to 304769)
 AUTHORS Read, T.D., Myers, G.S., Brunham, R.C., Nelson, W.C., Paulsen, I.T., Heidelberg, J., Holtzapple, E., Khouri, H., Fedorova, N.B., Carty, H.A., Unayam, L.A., Haft, D.H., Peterson, J., Beanan, M.J., White, O., Salzberg, S.L., Hsia, R.C., McClarty, G., Rank, R.G., Bavoil, P.M. and Fraser, C.M.
 TITLE Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae
 JOURNAL Nucleic Acids Res. 31 (8), 2134-2147 (2003)
 MEDLINE 22569155
 PUBMED 12682364
 REFERENCE 2 (bases 1 to 304769)
 AUTHORS Read, T., Myers, G., Brunham, R., Nelson, W., Paulsen, I., Heidelberg, J., Holtzapple, E., Khouri, H., Fedorova, N., Carty, H., Unayam, L., Haft, D., Peterson, J., Beanan, M., White, O., Salzberg, S., Hsia, R.-C., McClarty, G., Rank, R., Bavoil, P. and Fraser, C.
 TITLE Direct Submission
 JOURNAL Submitted (29-OCT-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 FEATURES
 source 1..304769
 /organism="Chlamydomophila caviae GPIC"
 /mol_type="genomic DNA"
 /isolate="GPIC"
 /db_xref="taxon:227941"
 57..1052
 /gene="hemB"
 /locus_tag="CCA00001"
 57..1052
 /gene="hemB"
 /locus_tag="CCA00001"
 /note="similar to GB:L16784, SP:P33395, and PID:290384; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="delta-aminolevulinic acid dehydratase"
 /protein_id="AAP04754.1"
 /db_xref="GI:29834117"
 /translation="MSSALILRRPRNRKTAARDIYAETALLPDQFICFFVKEGKN ICESIELPGYVMSIDLLKIELRSLGKAVILFPVPSDDKDAYGYSNPNKI LCKSIYEVKAPDLPCVLTIDALDPYTHGHGMDQGVNDSVRFIGNIATLHAE NGADIVAPSDMMQDRAHRAKLDQSGWTNTLILSVKYSYSLNFPDLAGSHLOS GDKRNYQNPKNVLELLECSLDEQEGADMIMKPAGLYLDVLRVRNNTLPLAAVQ VSGEYALIAADKMGWLKEMAFYESLIAIKRAGADMIIISVATLILEMLVSSKF"

gene
 complement (1068..2480)
 /gene="ndrA"
 /locus_tag="CCA00002"
 /note="similar to SP:Q56586, GB:L16784, SP:P33395, and PID:290384; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="NADH:ubiquinone oxidoreductase, subunit A"
 /protein_id="AAP04755.1"
 /db_xref="GI:29834118"
 /translation="MKIATIRGLDLSLQSPKBSGFLKRIDPALVSVDIRPYSALALK TEYQDDVISSGSPIAEYKFGVFTSPVSGIVKIRRGESRLDVIWKTPGQNL KEYSYDLSLRAELLEIKFKGLFALFKQRPDIPALPTQPDVFINLADNRPTP STEKOLAVFSSREGEVFNVGVRVAKLFGLCPLHIVSTRLALPKDLKSIHLHKI TGPYBSGSPSTHIHYIAPITNEKDIIVTLSPOEVLATIGHLPLKGRILNEQVVALAGS LKSSLRVYITTKANFESILLPQDISSDVLSLSDPLTGRCDRESLPCIGWRDSTI SVLPNPKTQAQNFRLGINKPHTHTYLSGFLKRKHTYNDPTNLHGETRPLIDTEI YDKMPLKIPVPLIKAVITKDFELACMLGFLFVCEPDEFPALPTFIDPSKTEMLTIKD ALKDYAKETGILNLEYEDKE"
 complement (2502..2936)
 /locus_tag="CCA00003"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAP04756.1"
 /db_xref="GI:29834119"
 /translation="MNSNAQKTIIDSIKOILTYLNIIDFSPFGSSSSSSEVDEYILM EKTQEKIQIDKRSQDILOQTGTREQMEVFANNPDNFSPEEWLALENIRSSDAYKK ETALNEVTKDLSGGETKTSNNKTKTSSKKNKKNIWPL"
 3048..5201
 /gene="greA"
 /locus_tag="CCA00004"
 /note="similar to GB:L10328, SP:P04846, GB:M12163, PID:146649, PID:290510, GB:U00096, and PID:1790093; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="transcription elongation factor GreA"
 /protein_id="AAP04757.1"
 /db_xref="GI:29834120"
 /translation="MDYLEKLQVLDIEQPSFFNWEYCFNDVVRGAEVLQILEKV KHSLAPLFGKIADTVLPWEMIPDGKEDRVQLVDVQNTNAKPFYDAADIVNRK YAGRNFNALRVGLRDGREFQYLSRFDLHLSGSEFVFGSGGVMGVS3FL QQKVLIEFGVMMKADISFETAFKSLPLEKDHFLSRFGDPDGFSGYAKHPSEVIE MLKLDLGPKTAKEDLDVLPPEAMNRWQATKTKKNCITTPTKIDYRYN PGSDSLITOLETSLSOIETDSAKLVEIYQIRDLHSELKKNPREIVIVLQKAFMLIKK NQSLQTDRLILLSDFLGESSVLDSFSLSPREEIVSVNNISIVVLQKAFMLIKK HSPWENVFMKIFLSTSPSLRELTFRVGEDSYRKKLEKLECAEHMPTVPEVFT WFFELGSHDDGIFKPNDRKIERLFEALVFMVHASTPQKELGKIYTFVLSQRYL AVRNIEGAPLSYKELILLSTKCSQSFSSDLSVLQSLAEVHPDLKQNTVVVEEDIL WTPPESTKWMKNKLSQSGKVMDNAKETEDARVLGDLRENSYKFALEKRLQBEI RVLSERNRARIITKDIVTDKVGCGKYSLEPENGNLITYLSILGPDADPDNVLISL KSKLAQEMLDKSVGAAVQGGKKYKISRISQSIWDE"
 5353..6543
 /locus_tag="CCA00005"
 /note="similar to SP:P04693, GB:L10328, SP:P04846, GB:M12163, PID:146649, PID:290510, GB:U00096, and PID:1790093; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="amino acid biosynthesis aminotransferase, class I"

```
/protein_id="AAP04758.1"
/db_xref="GI:29834121"
/translation="MSFNLQPTSPDSIIQLQKLFLEDEREKVNLIQVYDPSKA
YGFSSVRKAQFLFMEYDEMKYLIPIGLSFFNQEMKLVFGVNSFVFGAALGCT
GALHGAIFSMAYTESKVIPEOTWGNHVRIFAQOGLVLYKPYSSSESLLFDEM
LSVLKAAKPNLSVLQCCNPTGMDLNEDMWTLEAEVMEKHOLLFPDFAYLGFGRG
IKEDRRIFEIFVGNVFFVAACAKNFSLYGRVGVFAVSHKITLEDLDKISSCLREK
ARGEYSSPPRHGAKIVSTILSDASKKEMISELDTBSSLGKTRARLVOARMNHVGH
PDLISQKGFPGFPGFVQVLFURSEKGIYITSGARPNLNGITDKNDIHVSQFABA
YQUS"
6839.7861
/locus_tag="CCA00006"
/locus_tag="CCA00006"
/locus_tag="CCA00006"
/note="similar to GB:L10328, SP:P04846, GB:M12163,
PID:146649, PID:290510, GB:J00096, and PID:I790093;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAP04759.1"
/db_xref="GI:29834122"
/translation="MAGSVAYRRHRNRNLYVYLVSILVLSLPRPFVNTQKHV
SLYTRVFGSVDISYITQDVENIILKORIALEERLQAYRIANTYPPVFEIISP
YFRKLITSVIRYDPAHWGSCWINVGNKQIKNSPVLGKVLGLVDVGVQQRRI
RLITDVGQPSVIAVRGIIQAWLVKDOIQLTGRHLRSLDSYILEKDKYKVLQLEL
SSSIQCSDDNTLLRGTLSGGGGLWKDETLVLHGEFCPSDGLKGLVGLDITGLD
GVFPFGLFAEITKVCOPREGACSYKIEAKSLABDLNLSVLILPAMENPNDRPDI
FGLWLD"
complement(7821..10958)
/gene="recB"
/locus_tag="CCA00007"
/complement(7821..10958)
/gene="recB"
/locus_tag="CCA00007"
/note="identified by match to TIGR protein family HMM
TIGR01866"
/codon_start=1
/transl_table=1
/product="exodeoxyribonuclease V, beta subunit"
/protein_id="AAP04760.1"
/db_xref="GI:29834123"
/translation="MKPFDIFNPQTSIRKYFLFASAGTGKFTIEQIVLRLLGVS
SHVENILAVTNAATNELKRIQDNLQALCOIKSVLEDPSPKSLPYLXDTNSVKLL
YMOVNALATIDRAVITFNGFCNVLQOHPKMQTKNAALTHSOAVFHLLKYL
QDLWSVLFEQFLAAYKNSKHTAFLTDKLLSSYTLQTFEHLPTCTTYDALS
WHNSIRKIDIDIPEDFLOILOHGDGFKOPPIDEDLRLFVECLYASITSVQLFSF
SKIAETHPNKRLTRYKPSKAFSHIELMSLYTEQFCNIELIFNLQVQLKWH
YTQWLSDESILAEELLSRSEIEITLURKFLVLIDEFQDTRKQNNYRSTAOIMEAV
NDAPSGSLFTGDPKQIYEWNRADLALYKAKSFPKSSQLHLINNYRSTAOIMEAV
NILFCSPFLQIPSYEPYIEVPLCPQSSSEHFENSHAPIHFFSYDDISQAAWISHT
ASYLQATYHPLGRMILVSDAFLITHCSIPVAFSKNKSIFHETETYLTLWL
FALLHPNYSKIORVILASSPHNLSDTIEKELYSYFSLSYVDYLLATFYHF
MTLOGEALLTPQDLSQEMERLCAVLDIVSSHPQQLYLFYTSGRWEESLFS
SYSDDTLTKTHASGLYDVVFCPLDLSKNSKASEWIRBMVACTRAKKQLF
IPQNSNCSNLATNYRVHVSYSIEDLAQOLTKEHPRLFSLSITRSETDVP
YKIPPTFTLSAPAKQIFSFSSVLAFLNVDNVPAPVASFVLPGRKTIIF
HKILENSPNKIFSKIVSTVTHVKNLVEEIIISOKLVTFSSPLSFATESFS
LQDICPSKICEESFLSNQOLMQGVLDLFFEVNDKYVIDWKTSLGETSLEYSE
NLNLYIEQNLDYQALYTHAKRFLQQLFVLVSVMGFIPIRGMDFEGNGLCLPNQ
KISNPTITRKCPATH"
complement(10945..14010)
/gene="recC"
/locus_tag="CCA00008"
/complement(10945..14010)
/gene="recC"
/locus_tag="CCA00008"
/note="similar to GB:L25079, SP:P42822, and PID:529423;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="exodeoxyribonuclease V, gamma subunit"
/protein_id="AAP04761.1"
/db_xref="GI:29834124"
Query Match 18.3%; Score 395; DB 1; Length 304769;
Best Local Similarity 53.3%; Pred. No. 4.2e-76;
Matches 1115; Conservative 0; Mismatches 865; Indels 111; Gaps 9;
QY 1 ATAAATCTTTTAAAAACAGCTCGCATTAATATTATTAGTGAGAGCTTTTATTTT 60
Db 21625 ATAAATCTTTTAAAAAGAGCTCACATCTAATTAAGAAGATGTGAGCTTTT 21684
QY 61 ATAATAAACTAAAGATTTTATTATTTTGGTGTATTTTGGTAACTTATGGTTC 120
Db 21685 TTATAATTAAAGAAAGATTTTATTAGTTTATTTATTATTATGTTAATCTCT 21738
QY 121 AGTCTCTATAGCAAAACAGACGACACCTCCGCGAGATCTTTCTGCTCAAGATGGA 180
Db 21739 CGGCTCATAGATGAATCAAAAACATTTGCTCTCGCAGACTTATCTACTTAGTATGCA 21798
QY 181 GCGAGTGCAGCAATAAAGAGTCCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC 240
Db 21799 GCGAGCGGAGCAAAATCGTAGCTCAGAAGCTCAATCGATAACCGGAATTCGAGGCAATC 21858
QY 241 TAAAGAATCTTAAGAC-----CGATTCTGTAGAGCAATGGAGCATCTTGGTCTGCAGT 294
Db 21859 CGGCTCATCGACCTTCTGTGAAACCTGTAGACGATTCAGCTTTTGGAGCTCTGCTCG 21918
QY 295 GAATGCTCTCATGAGTCTGCGAGATGCTGGGTATTTGCTTCTAGTAAACAGCTCGTCTTC 354
Db 21919 GAAAGTTTAGCAAGTCTTTTCGATAAGAT----- 21948
QY 355 TACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACTACGCTCTCTCCACC 414
Db 21949 -----TTCTCTGTTCTTTTTCAGGGAAACGACTCTCTCA 21981
QY 415 CAGTTTGTATGATTAAGACTCAAGCGGCAACAGCTTAGATATCTTTTACCTCAAC 474
Db 21982 AACTTTTGTAAAGCTTAAGACGCAAGCAGAGTGCAGAACTCCGCTGCAGAGTGCAC 22041
QY 475 ATCACTAGCTGCATACAGCTGCTTTTGGTGAGCCCTCCAGAGTCTGTCTACTAATAATA 534
Db 22042 TACTTATGATCAGTTTCAAGACCGCTTTACAGAGCTGCAGAGTGTGTGAAACAGATGGA 22101
QY 535 GGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAACTAAGAAATGC 594
Db 22102 GCAATTAGCTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22161
QY 595 CGATGAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGAATACCA 654
Db 22162 GCGCGAAGAGTACGCTGGATACACTTAACCAAGTTGGGTGCTTATCTTACAGAGAACCA 22221
QY 655 AGCGATTTCTTGACTCTTTAGGTAACTGACTTCTTCTGAGCTCTTACAGGCTGCTCTTCT 714
Db 22222 GAAGCTTCTTGGAGCAATAAAGACGACCTGCTGCTGATGATGATGATGATGATGATGATGATGAT 22281
QY 715 CCAATCTGTAGCAAAACAATAACAGAGCTGAGCTTTTAAAGAGATGCAAGATAACCC 774
Db 22282 ACAAGTAGAAACCAATAAACAACACTGCTGAGGAGTTAAATTAACAGTTGAAGGA----- 22335
QY 775 AGTAGTCCCAGGAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACATGCTAC 834
Db 22336 -----AGCTGGGGTTAGCTATCTCTGTATGATGATGATGATGATGATGATGATGATGATGAT 22389
QY 835 AGCGACACAGATAGAGAAAGATGGAATCGATTGGGATGCAATTTTTCAGGAGCAGAA 894
Db 22390 AGGNACTCAGTTACTGATTAATAGCATGCTATATCGGAAGCTTATGCTGCGGGGAAAAA 22449
QY 895 CGCTAGTGGAGCTGTAGAAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 954
Db 22450 CAGTACCGGCTGTGGGCAAGCAGACAGGCAAAATAACAGCCCGCAATATATAGAGCTTC 22509
QY 955 TAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGA 1014
Db 22510 CAAACAAACTATTGCAAAATGCAAAAAGTCATAGAAGACGCTCTTAAACTTGTCTCAGA 22569
```

```
QY 1015 CTCTCAATCTTCTAAGAGCGGAACAATGGTAATACAGCGCTGAGAAAGATCTTAAABA 1074
Db 22570 TTCTCGGATCTCAAAAGCTGCTTTGAAAGAACAAACAGCGCAAAAGATATCTCAA 22629
QY 1075 TATCAAACTTGCAGATGGTTCTGATGTTCCAAATCCAGGAAC-----TACAGT 1122
Db 22630 TGTGAACCTAGTGGTGTAGTGATGCTATCGGTGGTCTGGAGCTCCTGGTAGTGT 22689
QY 1123 TGGAGGCTCAAGCAACAAGAAAGTAGTATGTTAGTAGTATTCGTGTTTCCANGCTGTTAGA 1182
Db 22690 GGGGACTTCTCAAAATCGCGGTGCTACTCTAGGGAAGTTCCGGTATCGATGTTATTGAC 22749
QY 1183 TGAATGCTGAATACAGACCGCTTCCATTTTGATGCTCGGTGTTTCGTCAGATGATTCACAT 1242
Db 22750 TGAATGTTAATGAACCCGAGGATCATATGCAAGGTTTCAGAAATGAATGATGATA 22809
QY 1243 GTTCAATACGGAATCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACACAAGCTAG 1302
Db 22810 CTTCCATGATCAAAACTCTGATTTTACAGCGCTTTTGAAGAGATTATGAATCAAGTAAC 22869
QY 1303 AG-----CAGCGAAACCGCTGGAGATGACAGTCTGCTGCAGCGCTGCGCATGC 1353
Db 22870 CGACTTATCAACGACGATCAATCTCGAGATCGGAAGCTTACAGCAACAATCAAGAAAT 22929
QY 1354 TCAGAAAGCTTTAGAAAGCGGCTCTAGTAAAGCTGGSCAACACAGGGCATATCTCAATGC 1413
Db 22930 ACAAACCATACAGATGCCCTTCAAGGATCGCGTCAAGACGGCATGATCAATGC 22989
QY 1414 TTTAGGACAGATCGCTTCTGCTGTTGTGAGCGCAGGAGTTCTTCCCTCCGCTCGACCAAG 1473
Db 22990 TTTAGGAGCTATAACACAGCAGCTTCAATTTCTACAGGAGCTCTATCGCTTCTGCAA 23049
QY 1474 TTCTATAGGCTCATCTGTAACAGCTTTTACAAGAC---CTCAAAATCTACAGTTCTGA 1530
Db 23050 TCAAGTGGATCAGCTTAAAGCAGCTTTTACAAACAGGATCTACTGCTCGGAGTTCTAA 23109
QY 1531 TTATAAAACACA---GATATCAGCAGGTTATGCTTACAAATCCATCAATGATGCCTA 1587
Db 23110 ATCTTACCGGATCTCTTATCTGAGGATGAGGGCATATCAATCTTTAATGATGTGA 23169
QY 1588 TGGTAGGCGCAGAAATGATCGACTCGTGTGATGATGATAAACAATGTAGTACCCCGCTCT 1647
Db 23170 CTCACG---TAGTAGTGCATCTAACCGTGAGTTTGTAGTACGTACGATCCAGCAT 23226
QY 1648 CACAGATCGTTCTTAGACACGACAGAGCTCGAGGACCGAAGAAACAGATCAAGC 1707
Db 23227 AACGACAGCTTTCTAGAACAGAACTCGGCTCGTGATAATGATAACGAGCTCAGCG 23286
QY 1708 CCTCGCTAGGGTGATTTCTGGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTC 1767
Db 23287 TTTTGAAGAACTATAGCTGTAATAGTATACTCTTGGGATGTTTATGCATCCGTAGG 23346
QY 1768 GGCACATCAATCTGTAATGAGATCATCCAGTGCATCTCAAGCGGAATTAATGAGAGAT 1827
Db 23347 TGTATTGCAACATCTGTAGGTGTTATCAAAATAATCCCAAGCGAATGAAGAAGAAAT 23406
QY 1828 CACACAAAAGCTTACATCGGACGTGACAAAGCTCCACAGTTTGGCTATCTCTTATGTGA 1887
Db 23407 CAAACAGAGCTCACCTTCTGAGGTTAGAAAGCTCCGAGTCAAGTTATCTCATGTACA 23466
QY 1888 ACTTCTTAATGACTCTACACAGAAGTTTCATAGCTTAATAGAAAGTTTGTGCTGAAGG 1947
Db 23467 GCCTTCTAACGACTCTACGAGAGAGTTTCAATGCTCACTCGAATGAATTTGTTCAAGG 23526
QY 1948 ATCTAGGACAGAGCTGAATAAAGCACTTCTCTTTTGAACGAACTCCCTGTTGTTATCA 2007
Db 23527 ATCGAAAGACTTTGCCGAGCAAAAGAGCTTGGCTTTTGAAGAACAGCTTTGTTCTATCCA 23586
QY 2008 GCAGGTCTGCTCAATCGGCTCTCTATATCTGTTTATCTCCATAACA 2058
Db 23587 GCAGGTTATTAGTGAACGTAGCACTCTGTTCTCGGGATACCTACAGTAAAA 23637
```

```
RESULT 15
BD264444
LOCUS
DEFINITION
Compounds and methods for treatment and diagnosis of chlamydial
infection.
ACCESSION
BD264444
VERSION
BD264444.1 GI:33074212
KEYWORDS
JP 2002531129-A/32.
SOURCE
Chlamydia trachomatis
ORGANISM
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE
1 (bases 1 to 1537)
AUTHORS
Probst, P., Bhatia, A., Skeiky, Y.A.W., Fling, S.P., Jen, S. and
Stromberg, E.J.
TITLE
Compounds and methods for treatment and diagnosis of chlamydial
infection
JOURNAL
Patent: JP 2002531129-A 32 24-SEP-2002;
CORIXA CORP
COMMENT
PN JP 2002531129-A/32
PD 24-SEP-2002 JP 2000586916
PF 08-DEC-1999 JP 2000586916
PR 08-DEC-1998 US 09/208277,08-APR-1999 US 09/288594 PR
01-OCT-1999 US 09/410568,22-OCT-1999 US 09/426571 PI PETER
PROBST, AJAY BHATIA, YASIR A W SKEIKY, STEVEN P FLING, PI SHYIAN
JEN, ERICA JEAN STROMBERG
PI ERICA JEAN STROMBERG
PC C12N15/09, C12N15/09, A61K35/12, A61K38/00, A61K39/118, A61K39/395,
PC A61K39/395,
PC A61K45/00, A61K48/00, A61P33/00, C07K14/295, C07K16/12, C07K19/00,
PC C12N1/15,
PC
C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12Q1/68, G01N33/53, G01N33/ PC
532,
PC
G01N33/569// (C12N1/19, C12R1:01), (C12N15/09, C12R1:01), (C12Q1/68, PC
C12R1:01),
PC C12N15/00, C12N15/00, C12N5/00, A61K37/02, (C12N15/00, C12R1:01) CC
Compounds and methods for treatment and diagnosis of CC
chlamydial infection
PH Key Location/Qualifiers
FT source 1..1537
/organism='Chlamydia trachomatis'.
FEATURES
Location/Qualifiers
source 1..1537
/organism='Chlamydia trachomatis'
/mol_type='genomic DNA'
/db_xref='taxon:813'
ORIGIN
Query Match 8.7%; Score 187.2; DB 6; Length 1537;
Best Local Similarity 53.0%; Pred. No. 2.7e-30;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;
QY 1167 TTTTCCATGCTGTGTAGATGATGCTGAAAATGAGACCGCTTCCATTTTGTATGCTCGGTTTC 1226
Db 186 TTTTCTTTGTTGCTGTGATGATGATGACAAATGGAATGCGAGCGATTGCAATGCAAGGTTTC 245
QY 1227 GTCAGATGATTCACATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCAACAGGAGC 1286
Db 246 GATCTATGATCGAACAAATTTAATGTAAACAATCTGCAACAGCTAAAGAGCTACAAGCTA 305
QY 1287 TCGCAGCAACAGCTAGAGCAGCGGAAGCGCTGGAGATGACAGTGTCTGTCGAGCGCTCG 1346
Db 306 TGGAGGCTCAGCTGA---CTGCGATGTCAGATCAACTGTTGTTGCGGATGCGAGCTCC 362
QY 1347 CAGATCTCAGAAAGCTTTAGAAAGCGGCTCTTAGTAAAGCTGGGCAACAAAC----- 1397
Db 363 CAGCCGAATACAGCAATCAAAAGATGCTCTTGGCAAGCTTTGAAACAACCAATCAGCAG 422
QY 1398 AGGGCATATCAATGTTTGTAGGACAGATGCTTCTGCTCTGTTGTGAGCGGAGGATTC 1457
```

Db 423 ATGCTTAGCTACAGCTATGGACAAGTGGCTTTTGACGCTGCCAAGGTTGGAGAGGCT 482
QY 1458 CTCCTCGCTGAGCAAGTCTTATAGGTCATCTGTAAAAAGAGCTTTTACAGACCTCAAAAT 1517
Db 483 CCGCAGCAACAGCTGGCACTGTCCAGATGAATGTAACACAGCTTTTACAGACAGCGTTT 542
QY 1518 CTACAGGTTCT-----GATTATAAAACACAGATATCAGCAGGTTATGATGCTTACAAAT 1571
Db 543 CTTGCACTTCTTCCAGCTCTTATGCGACGACCTTCCGATGGATATTCGCTTACAAA 602
QY 1572 CCATCAATGATGCTATGGTAGGCAAGAAATGATCGACTCGTGATGTGATATAAACAAATG 1631
Db 603 CACTGAACCTTTATATTCGAAAGCAGAAAGCGCGTGCAGTCAGCTATTAGTCAAACG 662
QY 1632 TAAGTACCCCGCTCTCACAGGATCGGTTCTTAGAGCAGCAAGAGCTCGAGGACCCAG 1691
Db 663 CAAATCCCGCTTTCCAGAAAGGTTTCTCGTTCTGGCATAGAAAGTCAAGGACGCGAGTG 722
QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTTTGGAGATG 1751
Db 723 CAGATGTAGCCAAAGAGCAGAGAACTATTGTGAGATAGCCAAAAGTTAGGTGATG 782
QY 1752 TCTATAGTCAAGTTTCGGCACTCAATCTGTAATGAGATCATCCAGTGAATCCTCAAG 1811
Db 783 TATATAGCCGCTTACAGGTTCTGGATTCTTTGATGTTAGATTGTGAGCAATCCGCAAG 842
QY 1812 CGAATAATGAGGAGATCAGAAAGCTTACATCGGCAAGTGACAAAGCCCTCCACAGTTTG 1871
Db 843 CAAATCAAGAAAGATATTGCAAGAGCTACGGCATCTATTAGCAAGCTCCCAATTTG 902
QY 1872 GCTATCCTTATGTGCAACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAATTAGAAA 1931
Db 903 GGTATCCTGCTGTTTCAAGATTCTGTGATAGCTTGAGAGTTTGTGACATTTGAAA 962
QY 1932 GTTTGTTTGTGAAGATCTAGGACAGAGCTGAAATATAAGCACTTTCCTTTGAAACGA 1991
Db 963 GAGAGTTTGTGTGAGGGAACGTAGTCTCGCAAGATCTCAAGAGAAATGCGTTTAGAAAAC 1022
QY 1992 ACTCCTTTGTTTATTTCAGCAGGCTGCTCAATATCGGCTCTCTATATTCTGGTTATCTCC 2051
Db 1023 AGCCCGTTTCAATCAACAGGTTGTTGTAACATTTGCTTCTCTATTCTCTGTTAICTTT 1082
QY 2052 AATAAC 2057
Db 1083 CTTAAC 1088

Search completed: March 24, 2004, 01:57:06
Job time : 8390.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:23:54 ; Search time 744.183 Seconds
(without alignments)
10716.954 Million cell updates/sec

Title: US-10-608-559-1
Perfect score: 2156
Sequence: 1 ataaatttttaaaacagg.....gatcttcaataataaaga 2156

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2156	100.0	2156	10	US-09-564-479-1
2	2143.4	99.4	1230025	15	US-10-289-762-1
3	1956	90.7	1956	9	US-09-841-132-385
4	1956	90.7	1956	10	US-09-841-260-63
5	1956	90.7	1956	13	US-10-007-693-63
6	1956	90.7	1956	15	US-10-312-273-14
7	1848.8	85.8	1852	10	US-09-564-479-3
8	1465.8	68.0	2238	10	US-09-564-479-7
9	1455	67.5	1456	10	US-09-564-479-5
10	187.2	8.7	1537	9	US-09-841-132-57
11	185.4	8.6	1941	9	US-09-841-132-412
12	184	8.5	1171	10	US-09-841-260-13
13	184	8.5	1171	13	US-10-007-693-13
14	184	8.5	1834	10	US-09-841-260-38
15	184	8.5	1834	13	US-10-007-693-38

16	183	8.5	1983	10	US-09-841-260-28	Sequence 28, Appl
17	183	8.5	1983	13	US-10-007-693-28	Sequence 28, Appl
18	170	7.9	1550	9	US-09-905-119-1	Sequence 1, Appl
19	49	2.3	1860	12	US-10-282-122A-37518	Sequence 37518, A
20	47	2.2	9373	12	US-10-221-613-178	Sequence 178, App
21	47	2.2	9373	14	US-10-311-455-1264	Sequence 1264, App
22	45.2	2.1	2404	15	US-10-104-047-1740	Sequence 1740, App
23	45.2	2.1	3439	15	US-10-151-927-53	Sequence 53, Appl
24	44.4	2.1	2499	9	US-09-842-552-92	Sequence 92, Appl
25	44.4	2.1	4997	12	US-10-282-122A-35506	Sequence 35506, A
26	44	2.0	3673778	14	US-10-312-841-2	Sequence 2, Appl
27	43.8	2.0	702	15	US-10-027-632-45880	Sequence 45880, A
28	43.8	2.0	702	15	US-10-027-632-45881	Sequence 45881, A
29	43.6	2.0	6341	14	US-10-311-455-1617	Sequence 1617, App
30	43	2.0	18218	14	US-10-311-455-1922	Sequence 1922, App
31	43	2.0	18585	14	US-10-240-485-161	Sequence 161, App
32	43	2.0	3673778	14	US-10-312-841-1	Sequence 1, Appl
33	42.8	2.0	302	12	US-10-424-599-63444	Sequence 63444, A
34	42.8	2.0	16688	14	US-10-311-455-293	Sequence 293, App
35	42.6	2.0	1431	12	US-10-424-599-127608	Sequence 127608, App
36	42.6	2.0	6126	14	US-10-311-455-1418	Sequence 1418, App
37	42.4	2.0	330	10	US-09-814-353-5359	Sequence 5359, App
38	42.4	2.0	330	10	US-09-814-353-11646	Sequence 11646, A
39	42.4	2.0	5503	16	US-10-257-166-104	Sequence 104, App
40	42.2	2.0	897	12	US-10-424-599-127609	Sequence 127609, App
41	42.2	2.0	6692	12	US-10-221-714A-131	Sequence 131, App
42	42	1.9	5535	14	US-10-311-455-909	Sequence 909, App
43	41.8	1.9	5278	12	US-10-221-714A-97	Sequence 97, Appl
44	41.8	1.9	5278	14	US-10-311-455-795	Sequence 795, App
45	41.8	1.9	6338	14	US-10-311-455-1099	Sequence 1099, App

ALIGNMENTS

RESULT 1

US-09-564-479-1
; Sequence 1, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OOMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(2053)
US-09-564-479-1

Query Match 100.0%; Score 2156; DB 10; Length 2156;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTTAAACAGGCTCGCATTAAATATTAGTAGAGAGCTTTTTTTTTTTTTTTT 60
Db 1 ATAAATCTTTTAAACAGGCTCGCATTAAATATTAGTAGAGAGCTTTTTTTTTTTTTTTT 60
QY 61 ATAAATAACTAAAGAGATTTTATTATTTTGTAGTTTATGTTATCCATTATGGTCC 120

Db 61 ATATAAATACTAAAGATTTTATTATTTTGGTTTATGGTTAAATCTTATGGTCC 120
QY 121 AGGTCTCTATAGACGAAACAGACGACACCTCCCGCAGATCTTCTGCTCAAGGATGGA 180
Db 121 AGGTCTCTATAGACGAAACAGACGACACCTCCCGCAGATCTTCTGCTCAAGGATGGA 180
QY 181 GCGAGTCGAGCAAAATAAGAGTGGGAAGCTCAAGAAATAGCAGGTGGGAAGCTAAGCC 240
Db 181 GCGAGTCGAGCAAAATAAGAGTGGGAAGCTCAAGAAATAGCAGGTGGGAAGCTAAGCC 240
QY 241 TAAAGATCTAGACCGATCTCTGATAGCGATGGAGCACTCTGCGCTTCTGAGTGAATGC 300
Db 241 TAAAGATCTAGACCGATCTCTGATAGCGATGGAGCACTCTGCGCTTCTGAGTGAATGC 300
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGCTTCTACTAG 360
Db 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGCTTCTACTAG 360
QY 361 CAGATCTGACAGCTGGAATCAACGACGACCGCAGCTAGCGCTTCTCCACCGATC 420
Db 361 CAGATCTGACAGCTGGAATCAACGACGACCGCAGCTAGCGCTTCTCCACCGATC 420
QY 421 TGAATGATTAAGACTCAAGCGCAAAACAGCTTACGATAGTATCTTTTACCTCAACATCACT 480
Db 421 TGAATGATTAAGACTCAAGCGCAAAACAGCTTACGATAGTATCTTTTACCTCAACATCACT 480
QY 481 AGCTGACATVACAGGCTGCTTTGGTAGCGCTTCCAGGATGCTGTCACTAAATATAAGGATAC 540
Db 481 AGCTGACATVACAGGCTGCTTTGGTAGCGCTTCCAGGATGCTGTCACTAAATATAAGGATAC 540
QY 541 AGCGGTACTGATGAGAAACCGCAATCGTGGGAGTGGGAACTAAGAAATGCCGATGC 600
Db 541 AGCGGTACTGATGAGAAACCGCAATCGTGGGAGTGGGAACTAAGAAATGCCGATGC 600
QY 601 AGTTAAAGTTGGCGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT 660
Db 601 AGTTAAAGTTGGCGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT 660
QY 661 TCTTGACTCTTTAGGTAAATGACTCTTCTGACCTCTTACAGGCTGCTTCTTCCCAATC 720
Db 661 TCTTGACTCTTTAGGTAAATGACTCTTCTGACCTCTTACAGGCTGCTTCTTCCCAATC 720
QY 721 TGTAGCAAAATAACAAAGCAGCTGAGCTCTTAAAGAGTGCAGATAACCCAGTAGT 780
Db 721 TGTAGCAAAATAACAAAGCAGCTGAGCTCTTAAAGAGTGCAGATAACCCAGTAGT 780
QY 781 CCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db 781 CCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
QY 841 ACAGATAGAGAAGATGGAATGCGATTAGGATGATATTTTGCAGGACAGAACGCTAG 900
Db 841 ACAGATAGAGAAGATGGAATGCGATTAGGATGATATTTTGCAGGACAGAACGCTAG 900
QY 901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGCAACATAGATTCAGCTTAAAGC 960
Db 901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGCAACATAGATTCAGCTTAAAGC 960
QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
Db 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
QY 1021 AATTCCTCAAGAACGGAACAAATGTTAATACAGGCTGAGAAAGATCTTAAATAATCAA 1080
Db 1021 AATTCCTCAAGAACGGAACAAATGTTAATACAGGCTGAGAAAGATCTTAAATAATCAA 1080
QY 1081 ACCTGCAGATGTTCTGATGCTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGCAACA 1140
Db 1081 ACCTGCAGATGTTCTGATGCTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGCAACA 1140
QY 1141 AGGAAGTAGTATTGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAATGAGAC 1200

Db 1141 AGGAAGTAGTATTGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC 1200
QY 1201 CGCTTCCAAATTTGATGCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATPCC 1260
Db 1201 CGCTTCCAAATTTGATGCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATPCC 1260
QY 1261 TGATTTCTAGCTGCCCAACAGGAGCTCGCAGCAAGCTAGAGCAGGGAAGCGCTGG 1320
Db 1261 TGATTTCTAGCTGCCCAACAGGAGCTCGCAGCAAGCTAGAGCAGGGAAGCGCTGG 1320
QY 1321 AGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTCTAGG 1380
Db 1321 AGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTCTAGG 1380
QY 1381 TAAAGCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Db 1381 TAAAGCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
QY 1441 TGTGAGCGCAGAGTTCTCCCGCTGCGAGCAAGTTCTATAGGTCATCTGTAAACAGCT 1500
Db 1441 TGTGAGCGCAGAGTTCTCCCGCTGCGAGCAAGTTCTATAGGTCATCTGTAAACAGCT 1500
QY 1501 TTACAAGACCTCAAAATCTACAGGTTCTGATTTATAAAAACAGATATCAGCAGGTTATGA 1560
Db 1501 TTACAAGACCTCAAAATCTACAGGTTCTGATTTATAAAAACAGATATCAGCAGGTTATGA 1560
QY 1561 TGCTTCAAAATCCATCAATGATGCTTATGTTAGGCAAGAAATGATGCGATCTGCTGATGT 1620
Db 1561 TGCTTCAAAATCCATCAATGATGCTTATGTTAGGCAAGAAATGATGCGATCTGCTGATGT 1620
QY 1621 GATTAACAAATGATGATGCTTCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGC 1680
Db 1621 GATTAACAAATGATGATGCTTCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGC 1680
QY 1681 TCGAGGACAGAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGCAATAGCAGAAC 1740
Db 1681 TCGAGGACAGAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGCAATAGCAGAAC 1740
QY 1741 TCTTGGAGATGCTATAGTCAAGTTTGGGCTACAAATCTGTAATGAGATCATCCAGTC 1800
Db 1741 TCTTGGAGATGCTATAGTCAAGTTTGGGCTACAAATCTGTAATGAGATCATCCAGTC 1800
QY 1801 GAATCCTCAAGCGAATAATGAGGAGATCAGAAAAGCTTACATCGGAGTGCAAGGCC 1860
Db 1801 GAATCCTCAAGCGAATAATGAGGAGATCAGAAAAGCTTACATCGGAGTGCAAGGCC 1860
QY 1861 TCCACAGTTTGGCTATCTTATGTCAACTTTCTAATGACTCTACACAGAAATTCATAGC 1920
Db 1861 TCCACAGTTTGGCTATCTTATGTCAACTTTCTAATGACTCTACACAGAAATTCATAGC 1920
QY 1921 TAAATTAGAAAGTTTGGTGGAGGATCTAGGACAGCAGCTGAAATTAAGAGCACTTTC 1980
Db 1921 TAAATTAGAAAGTTTGGTGGAGGATCTAGGACAGCAGCTGAAATTAAGAGCACTTTC 1980
QY 1981 CTTTGAACGAACTCTTGTATTATTCAGCAGCTGCTGCTCAATATCGGCTCTCTATATTC 2040
Db 1981 CTTTGAACGAACTCTTGTATTATTCAGCAGCTGCTGCTCAATATCGGCTCTCTATATTC 2040
QY 2041 TGGTTATCTCCAAATAACACCTTAAGTGTGTTTGGAGAGATTTATGTCGTTTGGT 2100
Db 2041 TGGTTATCTCCAAATAACACCTTAAGTGTGTTTGGAGAGATTTATGTCGTTTGGT 2100
QY 2101 AAGGCTTTGTTGGGCTTTACCAACACACTAGAACGATCTTCATTAATAAAGA 2156
Db 2101 AAGGCTTTGTTGGGCTTTACCAACACACTAGAACGATCTTCATTAATAAAGA 2156

RESULT 2

US-10-289-762-1/c
; Sequence 1, Application US/10289762
; Publication NO. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.

;; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
;; TITLE OF INVENTION: and treatment of infection
;; FILE REFERENCE: 9710-003-999
;; CURRENT APPLICATION NUMBER: US/10/289,762
;; CURRENT FILING DATE: 2003-03-27
;; NUMBER OF SEQ ID NOS: 6849
;; SEQ ID NO 1
;; LENGTH: 1230025
;; TYPE: DNA
;; ORGANISM: Chlamydia pneumoniae
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)-(15000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (15001)-(30000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (30001)-(45000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (45001)-(60000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (60001)-(75000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (75001)-(90000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (90001)-(105000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (105001)-(120000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (120001)-(135000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (135001)-(150000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (150001)-(165000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (165001)-(180000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (180001)-(195000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (195001)-(210000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (210001)-(225000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (225001)-(240000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (240001)-(255000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (255001)-(270000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (270001)-(285000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (285001)-(300000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (300001)-(315000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (315001)-(330000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (330001)-(345000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (345001)-(360000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (360001)-(375000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (375001)-(390000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (390001)-(405000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (405001)-(420000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (420001)-(435000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (435001)-(450000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (450001)-(465000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (465001)-(480000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (480001)-(495000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (495001)-(510000)

/	OTHER INFORMATION: n=a or c or g or t	Db	828458	TAAAGAACTTAAGACCGATTCTGTAGACGATGGAGCATCTTGGCTTCGCAATGATGC	828399
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	301	TCTCATGTGCTGGCAGATAAGCTGGGTATTCCTCTTAGTAAACAGCTGCTCTTACTAG	360
/	LOCATION: (510001)..(525000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828398	TCTCATGTGCTGGCAGATAAGCTGGGTATTCCTCTTAGTAAACAGCTGCTCTTACTAG	828339
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	361	CAGATCTGCAGACGTGGACTCAACGACGACGCCACCTACGCTCCCTCCACCCACGTT	420
/	LOCATION: (525001)..(540000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828338	CAGATCTGCAGACGTGGACTCAACGACGACGCCACCTACGCTCCCTCCACCCACGTT	828279
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	421	TGATGATTATTAAGACTCAAGCGCAACAGCTTACGATACCTATCTTTACCTCAACATCACT	480
/	LOCATION: (540001)..(555000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828278	TGATGATTATTAAGACTCAAGCGCAACAGCTTACGATACCTATCTTTACCTCAACATCACT	828219
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	481	AGCTGACATACAGGCTGCTTGGTGAGGCTCCAGGATGCTGTCACTAATATAAAGATAC	540
/	LOCATION: (555001)..(570000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828218	AGCTGACATACAGGCTGCTTGGTGAGGCTCCAGGATGCTGTCACTAATATAAAGATAC	828159
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	541	AGCGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC	600
/	LOCATION: (570001)..(585000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828158	AGCGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC	828099
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	601	AGTTAAAGTTGGCGCGCAAAATTACAGAATTACGAAATATGCTTCGGATAACCAAGCGAT	660
/	LOCATION: (585001)..(600000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828098	AGTTAAAGTTGGCGCGCAAAATTACAGAATTACGAAATATGCTTCGGATAACCAAGCGAT	828039
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	661	TCTTGACTCTTTAGGTAACTGACTCTCTTGACCTCTTACAGGCTGCTCTTCTCAATC	720
/	LOCATION: (600001)..(615000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828038	TCTTGACTCTTTAGGTAACTGACTCTCTTGACCTCTTACAGGCTGCTCTTCTCAATC	827979
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	721	TGTAGCAACAAATAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	780
/	LOCATION: (615001)..(630000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827978	TGTAGCAACAAATAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	827919
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	781	CCCAGGAAACCGCTGCAATTGCTCAATCTTTAGTTGATCAGAAGATGCTTACAGCGAC	840
/	LOCATION: (630001)..(645000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827918	CCCAGGAAACCGCTGCAATTGCTCAATCTTTAGTTGATCAGAAGATGCTTACAGCGAC	827859
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	841	ACAGATAGAGAAAGATGCAAAATGCGATTAGGAGTGCATATTTTGACGACGACGCTAG	900
/	LOCATION: (645001)..(660000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827858	ACAGATAGAGAAAGATGCAAAATGCGATTAGGAGTGCATATTTTGACGACGACGCTAG	827799
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	901	TGGAGCTGTAGAAAATGTCTAAATCTAATAACAGTATAAGCAACATAGATTACGCTAAAGC	960
/	LOCATION: (660001)..(675000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827798	TGGAGCTGTAGAAAATGTCTAAATCTAATAACAGTATAAGCAACATAGATTACGCTAAAGC	827739
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	961	AGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGACTCTCC	1020
/	LOCATION: (675001)..(690000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827738	AGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGACTCTCC	827679
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1021	AATTCCTCAAGAGCGGAAACAAATGCGTAATACAGGCTCAGAAAAGTCTTAAATAATCAA	1080
/	LOCATION: (690001)..(705000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827678	AATTCCTCAAGAGCGGAAACAAATGCGTAATACAGGCTCAGAAAAGTCTTAAATAATCAA	827619
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1081	ACCTGACAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGCTCCAAACACA	1140
/	LOCATION: (705001)..(720000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827618	ACCTGACAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGCTCCAAACACA	827559
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1141	AGGAAGTAGTATGTTAGTATTCGTTGTTTCAATGCTGTTAGATGATGCTGAAAATGAGAC	1200
/	LOCATION: (720001)..(735000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827558	AGGAAGTAGTATGTTAGTATTCGTTGTTTCAATGCTGTTAGATGATGCTGAAAATGAGAC	827499
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1201	CGCTTCCATTTTGTGTTGCTGTTTTCGTCAGATGATTCACATGTTCAATACGGAATAATCC	1260
/	LOCATION: (735001)..(750000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827498	CGCTTCCATTTTGTGTTGCTGTTTTCGTCAGATGATTCACATGTTCAATACGGAATAATCC	827439
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1261	TGATTCCTCAAGTGCCTCAACAGGAGCTCGACACAAAGCTAGACGAGCAAGCGCTGG	1320
/	LOCATION: (750001)..(765000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827438	TGATTCCTCAAGTGCCTCAACAGGAGCTCGACACAAAGCTAGACGAGCAAGCGCTGG	827379
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1321	AGATGACAGTCTGCTGACGCTGGCAGATGCTCAGAAAAGCTTTAGAGGGCTCTAGG	1380
/	LOCATION: (765001)..(780000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827378	AGATGACAGTCTGCTGACGCTGGCAGATGCTCAGAAAAGCTTTAGAGGGCTCTAGG	827319
/	FEATURE:				

Query Match 99.4%; Score 2143.4; DB 15; Length 1230025;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2155; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT 60
 Db 828698 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT 828639

QY 61 ATAATTAARACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
 Db 828638 ATAATTAARACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 828579

QY 121 AGTCTCTATAGCAACAGGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA 180
 Db 828578 AGTCTCTATAGCAACAGGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA 828519

QY 181 GCGAGTGCAGCAATAAGAGTGGGAAGCTCAAGAAATAGCAGGTGCGGAAGCTTAAGCC 240
 Db 828518 GCGAGTGCAGCAATAAGAGTGGGAAGCTCAAGAAATAGCAGGTGCGGAAGCTTAAGCC 828459

QY 241 TAAAGATCTAAGACCGATCTCTAGAGGATGGAGCATCTTGGCTTCTGCAATGATGC 300

```
1381 TAAAGCTGGGCAACAAACAGGCGATCAATGCTTTAGACAGATCGCTTCTGCTGCTG- 1439
Db |||||||
827318 TAAAGCTGGGCAACAAACAGGCGATCAATGCTTTAGACAGATCGCTTCTGCTGCTG- 827259
Qy |||||||
1440 TTGTGAGCGCAGGAGTCTCTCCGCTGCGAGCAAGTTCTATAGGCTCATCTGTAAACAGC 1499
Db |||||||
827258 TTGTGAGCGCAGGAGTCTCTCCGCTGCGAGCAAGTTCTATAGGCTCATCTGTAAACAGC 827199
Qy |||||||
1500 TTACAGACCTCAAAATCTACAGCTTCTGATTAATAAACAACAGATATCAGCAGGTTATG 1559
Db |||||||
827198 TTACAGACCTCAAAATCTACAGCTTCTGATTAATAAACAACAGATATCAGCAGGTTATG 827139
Qy |||||||
1560 ATGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGAAATGATCGATCTGTGATG 1619
Db |||||||
827138 ATGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGAAATGATCGATCTGTGATG 827079
Qy |||||||
1620 TGATAAACAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGACAGCAACAGAAAG 1679
Db |||||||
827078 TGATAAACAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGACAGCAACAGAAAG 827019
Qy |||||||
1680 CTCGAGGACCAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGCAATAGCAGAA 1739
Db |||||||
827018 CTCGAGGACCAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGCAATAGCAGAA 826959
Qy |||||||
1740 CTCCTGGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGATGAGATCATCCAGT 1799
Db |||||||
826958 CTCCTGGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGATGAGATCATCCAGT 826899
Qy |||||||
1800 CGAATCTCAAGCAATATAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGC 1859
Db |||||||
826898 CGAATCTCAAGCAATATAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGC 826839
Qy |||||||
1860 CTCGACGTTTGGCTATCTTATGTCGAATCTTCAATGACTCTACAGAGTTTCATG 1919
Db |||||||
826838 CTCGACGTTTGGCTATCTTATGTCGAATCTTCAATGACTCTACAGAGTTTCATG 826779
Qy |||||||
1920 CTAAATTAGAAAGTTGTTTGTGTAAGGATCTAGACAGCAGCTGAATATAAAGCACTTT 1979
Db |||||||
826778 CTAAATTAGAAAGTTGTTTGTGTAAGGATCTAGACAGCAGCTGAATATAAAGCACTTT 826719
Qy |||||||
1980 CTTTGTAAACGAATCTCTGTTTATTCAGCAGGCTGCTGCTCAATATGCGCTCTCATATT 2039
Db |||||||
826718 CTTTGTAAACGAATCTCTGTTTATTCAGCAGGCTGCTGCTCAATATGCGCTCTCATATT 826659
Qy |||||||
2040 CTGGTTATCTCAATAACACACCTAAGTTCGTTTGGAGAGATTTATATGCTTTGG 2099
Db |||||||
826658 CTGGTTATCTCAATAACACACCTAAGTTCGTTTGGAGAGATTTATATGCTTTGG 826599
Qy |||||||
2100 TAAGGCTTTGTTGAGGCTTACCAACACATAGAACGATCTTCAATATAAAGA 2156
Db |||||||
826598 TAAGGCTTTGTTGAGGCTTACCAACACATAGAACGATCTTCAATATAAAGA 826542

RESULT 3
US-09-841-132-385
; Sequence 385, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.46283
; CURRENT APPLICATION NUMBER: US/09841132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq
; SEQ ID NO 385
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
```

1121 GTTGGAGGCTCAAGCAACAGGAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTTA 1180
1021 GTTGGAGGCTCAAGCAACAGGAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTTA 1080
1181 GATGATGCTGAAAATGAGACCGCTTCCATTTTGAATCTGTTGGTTCCTGAGATGATTTAC 1240
1081 GATGATGCTGAAAATGAGACCGCTTCCATTTTGAATCTGTTGGTTCCTGAGATGATTTAC 1140
1241 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGAGCTCCGACGCAAGCT 1300
1141 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGAGCTCCGACGCAAGCT 1200
1301 AGAGCAGCAAGAACCGCTGAGATGACAGTCTGCTGACGCGTGGCAGATGCTCAGAAA 1360
1201 AGAGCAGCAAGAACCGCTGAGATGACAGTCTGCTGACGCGTGGCAGATGCTCAGAAA 1260
1361 GCTTTAGAAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGA 1420
1261 GCTTTAGAAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGA 1320
1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGAGTTCTCTCCGCTGCAGCAAGTCTCTATA 1480
1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGAGTTCTCTCCGCTGCAGCAAGTCTCTATA 1380
1481 GGGTCATCTGTAACACAGCTTTACAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1540
1381 GGGTCATCTGTAACACAGCTTTACAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1440
1541 CAGATATCAGCAGGTTATGATGCTTACAATCCATCAATGATGCCCTATGGTAGGCGACA 1600
1441 CAGATATCAGCAGGTTATGATGCTTACAATCCATCAATGATGCCCTATGGTAGGCGACA 1500
1601 AATGATGCGACTCGTCTGATGTAACAAATGTAAGTACCCCGCTCTCACAGATCCGTT 1660
1501 AATGATGCGACTCGTCTGATGTAACAAATGTAAGTACCCCGCTCTCACAGATCCGTT 1560
1661 CCTAGAGCAGAACAGAGCTCGAGGACCAAGAAAACAGATCAAGCCCTCGCTAGGGTG 1720
1561 CCTAGAGCAGAACAGAGCTCGAGGACCAAGAAAACAGATCAAGCCCTCGCTAGGGTG 1620
1721 ATTCTGGCAATAGCAAGACTTTGGAGATGCTATAGTCAAGTTTCGGCACTACAATCT 1780
1621 ATTCTGGCAATAGCAAGACTTTGGAGATGCTATAGTCAAGTTTCGGCACTACAATCT 1680
1781 GTAATCAGATCATCCAGTCGATCTCTCAAGCGAATAATGAGGAGATCAGCAAAAAGCTT 1840
1681 GTAATCAGATCATCCAGTCGATCTCTCAAGCGAATAATGAGGAGATCAGCAAAAAGCTT 1740
1841 ACATCGGCAAGTGAACAAAGCTCCACAGTTTGGCTATCCCTTATGTGCAACTTTCTAATGAC 1900
1741 ACATCGGCAAGTGAACAAAGCTCCACAGTTTGGCTATCCCTTATGTGCAACTTTCTAATGAC 1800
1901 TCTACACAGAGTTCTAGCTAATTTAGAAAGTTTCTTGGCTGAAGAGTCTAGGACAGA 1960
1801 TCTACACAGAGTTCTAGCTAATTTAGAAAGTTTCTTGGCTGAAGAGTCTAGGACAGA 1860
1961 GCTGAATAAAGCACTTCTTCTTTGAAAACGAATCTTGTGTTTATCAGCAGAGTGTCTG 2020
1861 GCTGAATAAAGCACTTCTTCTTTGAAAACGAATCTTGTGTTTATCAGCAGAGTGTCTG 1920
2021 AATATCGGCTCTCTATATTCTGTTATCTGTTATCTCCATAA 2056
1921 AATATCGGCTCTCTATATTCTGTTATCTCTCCATAA 1956

RESULT 4

US-09-841-260-63
; Sequence 63, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter

APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; FILE REFERENCE: 21(121.515
; CURRENT APPLICATION NUMBER: US/09/441.260
; CURRENT FILING DATE: 2004-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-63

Query Match 90.7%; Score 1956; DB 10; Length 1956;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 160
DB 1 ATGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 60
QY 161 CTTTCTGCTCAAGGATTTGGAGCGAGTGCAGCAATATAGAGTCCGGAAGCTCAAGAATA 220
DB 61 CTTTCTGCTCAAGGATTTGGAGCGAGTGCAGCAATATAGAGTCCGGAAGCTCAAGAATA 120
QY 221 GCAGTGGGGAAGCTTAAGCTTAAGAACTTAAGACCGGATTTCTGTAGAGCGATGGAGCATC 280
DB 121 GCAGTGGGGAAGCTTAAGCTTAAGAACTTAAGACCGGATTTCTGTAGAGCGATGGAGCATC 180
QY 281 TTGCTTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATTAAGCTGGGTATTTGCTTCTAGT 340
DB 181 TTGCTTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATTAAGCTGGGTATTTGCTTCTAGT 240
QY 341 AACAGCTGCTCTTCTACTAGCAGATCTGCAGACGTTGAGTCAACGACGCGACCGACCT 400
DB 241 AACAGCTGCTCTTCTACTAGCAGATCTGCAGACGTTGAGTCAACGACGCGACCGACCT 300
QY 401 AGCGCTCTCTCCACCCACCGTTTGATGATTATAGACTCAAGCGCAACACAGCTTACGATAT 460
DB 301 AGCGCTCTCTCCACCCACCGTTTGATGATTATAGACTCAAGCGCAACACAGCTTACGATAT 360
QY 461 ATCTTTAATCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 520
DB 361 ATCTTTAATCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 420
QY 521 GTCACATAATATAAGATACAGCGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGG 580
DB 421 GTCACATAATATAAGATACAGCGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGG 480
QY 581 GAAACTAAGAAATGCCGATGCACTTAAAGTTGGCGCGCAAAATTAACGAATTAAGCAATAT 640
DB 481 GAAACTAAGAAATGCCGATGCACTTAAAGTTGGCGCGCAAAATTAACGAATTAAGCAATAT 540
QY 641 GCTTCGGATAACCAAGGATTTCTGACTTTTGGTAACTGACTTCTTCGACCTCTTA 700
DB 541 GCTTCGGATAACCAAGGATTTCTGACTTTTGGTAACTGACTTCTTCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCTTCCATCTCTAGCAACAAATAACAAAGCAGCTGAGCTTCTTAAAGAG 760
DB 601 CAGGCTGCTCTTCTTCCATCTCTAGCAACAAATAACAAAGCAGCTGAGCTTCTTAAAGAG 560
QY 761 ATGCAAGATAACCCAGTAGTCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTAGT 820
DB 661 ATGCAAGATAACCCAGTAGTCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTAGT 720
QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATTCGATTAAGGATGCATAT 880
DB 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATTCGATTAAGGATGCATAT 780
QY 881 TTTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAAATGCTAAATCTTAATAACAGTTAAGC 940
DB 781 TTTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAAATGCTAAATCTTAATAACAGTTAAGC 840

941 AACATAGATTACGCTAAAGCAGCAATCGCTACTGTCTAAGACAAATAGCTGAAGCTCAG 1000
1001 AAAAGTTCCTCCGACCTCTCAATTCCTTCAAGAGCGGAAACAAATGGTAATAACAGCTGAG 1060
901 AAAAGTTCCTCCGACCTCTCAATTCCTTCAAGAGCGGAAACAAATGGTAATAACAGCTGAG 960
1061 AAAGATCTTAAATAATCAAACTGAGATGCTTCTGATGTTCCCAATCCAGGAATCA 1120
961 AAAGATCTTAAATAATCAAACTGAGATGCTTCTGATGTTCCCAATCCAGGAATCA 1020
1121 GTTGAGGCTCCAAGCAACAGGAAGTAGTATTGGTAGTATTGGTTCCTATGCTGTTA 1180
1021 GTTGAGGCTCCAAGCAACAGGAAGTAGTATTGGTAGTATTGGTTCCTATGCTGTTA 1080
1181 GATGATGCTGAAAATGAGACCGCTTCCATTTGATGCTGGGTTTCGTGAGATGATTCAC 1240
1081 GATGATGCTGAAAATGAGACCGCTTCCATTTGATGCTGGGTTTCGTGAGATGATTCAC 1140
1241 ATGTTCAATACGGAATCTGATTTCTCAAGCTGCCCAACAGAGCTCCAGCAACAGCT 1300
1141 ATGTTCAATACGGAATCTGATTTCTCAAGCTGCCCAACAGAGCTCCAGCAACAGCT 1200
1301 AGAGCAGCAAGCGCTCGAGATGACAGTCTGCTGACGCTGSCAGATGCTCAGAAA 1360
1201 AGAGCAGCAAGCGCTCGAGATGACAGTCTGCTGACGCTGSCAGATGCTCAGAAA 1260
1361 GCTTTAGAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGGATCTCAATGCTTTAGGA 1420
1261 GCTTTAGAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGGATCTCAATGCTTTAGGA 1320
1421 CAGATCGCTTCTGCTGCTGTGAGCGCAGAGTTCTCCCGCTGCAGCAAGTTCTATA 1480
1321 CAGATCGCTTCTGCTGCTGTGAGCGCAGAGTTCTCCCGCTGCAGCAAGTTCTATA 1380
1481 GGGTCATCTGTAACAAAGCTTTAAGACCTCAAAATCTACAGGTTCTGATTTAAACA 1540
1381 GGGTCATCTGTAACAAAGCTTTAAGACCTCAAAATCTACAGGTTCTGATTTAAACA 1440
1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTGATGAGGACGA 1600
1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTGATGAGGACGA 1500
1601 AATGATGCGACTCGTGATGATAAACAATGTAAGTACCCCGCTCTCACAGCTCCGTT 1660
1501 AATGATGCGACTCGTGATGATAAACAATGTAAGTACCCCGCTCTCACAGCTCCGTT 1560
1661 CTTAGAGCAGAACAGAGCTCGAGGCCAGAAAAACAGATCAAGCCCTCGCTAGGGTG 1720
1561 CTTAGAGCAGAACAGAGCTCGAGGCCAGAAAAACAGATCAAGCCCTCGCTAGGGTG 1620
1721 ATTTCCTGCAATAGCAGAACTCTTGAGATGCTATAGTCAAGTTTCGGCATCAATCT 1780
1621 ATTTCCTGCAATAGCAGAACTCTTGAGATGCTATAGTCAAGTTTCGGCATCAATCT 1680
1781 GTAATGCGAGATCATCCAGTCAATCTCAAGCGAATTAATGAGGATCAGACAAAGCTT 1840
1681 GTAATGCGAGATCATCCAGTCAATCTCAAGCGAATTAATGAGGATCAGACAAAGCTT 1740
1841 ACATGCGAGTGAACAGCTCCAGGTTTGGCTATTCCTTATGTCGAACTTTCTAATGAC 1900
1741 ACATGCGAGTGAACAGCTCCAGGTTTGGCTATTCCTTATGTCGAACTTTCTAATGAC 1800
1901 TCTACACAGAGTTCACTAGTAAATTAGAAAGTTTGTGTTGCTGAAGGATCTAGGACAGCA 1960
1801 TCTACACAGAGTTCACTAGTAAATTAGAAAGTTTGTGTTGCTGAAGGATCTAGGACAGCA 1860
1961 GCTGAAATAAAGCACTTTCTTTGAAACGAACTCTTCTTTTATTCAGCAGGTTGCTGCTC 2020
1861 GCTGAAATAAAGCACTTTCTTTGAAACGAACTCTTCTTTTATTCAGCAGGTTGCTGCTC 1920
2021 AATATCGGCTCTCTATATTCTGGTTATCTCCAAATA 2056

Db 1921 AATATCGGCTCTCTATATTCTGGTTATCTCCAAATA 1956

RESULT 5

US-10-007-693-63
; Sequence 63, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210421-545C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING/DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-63

Query Match 90.7%; Score 1956; DB 13; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 60
QY 161 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATAAAGAGTGCAGAAAGCTCAAAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATAAAGAGTGCAGAAAGCTCAAAAGAATA 120
QY 221 GCAGTGGGAAAGCTTAAGCTTAAAGATCTAAGACCGATTCTGTAGAGCGATGAGAGATC 280
Db 121 GCAGTGGGAAAGCTTAAAGCTTAAAGATCTAAGACCGATTCTGTAGAGCGATGAGAGATC 180
QY 281 TTGGTTCTGCGAGTGAATGCTCTCATGAGTCTGCGAGATAAGCTGGGTATGCTTCTAGT 340
Db 181 TTGGTTCTGCGAGTGAATGCTCTCATGAGTCTGCGAGATAAGCTGGGTATGCTTCTAGT 240
QY 341 AACAGCTCGTCTTCTACTAGCAGATCTCGACGCTGGAAGTCAACAGACGCGACCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTCGACGCTGGAAGTCAACAGACGCGACCGACCT 300
QY 401 AGCGCTCTCCACCCACCGCTTTGATGATTAAGACTCAAGCGCAAAACAGCTTACGATACT 460
Db 301 AGCGCTCTCCACCCACCGCTTTGATGATTAAGACTCAAGCGCAAAACAGCTTACGATACT 360
QY 461 ATCTTTACTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 520
Db 361 ATCTTTACTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 420
QY 521 GTCACATAATAAAGGATACAGCGCTACTGATCAGGAAACCGCAATCGCTGCGAGTGG 580
Db 421 GTCACATAATAAAGGATACAGCGCTACTGATCAGGAAACCGCAATCGCTGCGAGTGG 480
QY 581 GAACTAAGATGCGGATGCGATTAAGTTGGCGCGCAAAATACAGAAATGACGAAATAT 640
Db 481 GAACTAAGATGCGGATGCGATTAAGTTGGCGCGCAAAATACAGAAATGACGAAATAT 540
QY 641 GCTTCGGATAACCAAGGATTTGACTTTTAGTAACTGACTTCCTTCGACCTCTTA 700
Db 541 GCTTCGGATAACCAAGGATTTGACTTTTAGTAACTGACTTCCTTCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCTCCAATCTGTAGCAAAACAAATACAAAGCAGCTGAGCTTCTTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAATCTGTAGCAAAACAAATACAAAGCAGCTGAGCTTCTTAAAGAG 660
QY 761 ATGCMAGATAACCAAGGATGCTCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTCAT 820

Db 661 ATGCAAGATAACCCAGTAGTCCAGGGAAAAACGCGCTGCAATTGCTCAATCTTTAGTTGAT 720
Qy 821 CAGACAGATGCTACAGCAGACAGATAGAGAAAGATGGAATGCGATTAGGAGTGCATAT 880
Db 721 CAGACAGATGCTACAGCAGACAGATAGAGAAAGATGGAATGCGATTAGGAGTGCATAT 780
Qy 881 TTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 940
Db 781 TTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 840
Qy 941 AACATAGATTACAGTAAAGAGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTACAGTAAAGAGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAG 900
Qy 1001 AAAAGATTCCCGGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAAATACAGCTGAG 1060
Db 901 AAAAGATTCCCGGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAAATACAGCTGAG 960
Qy 1061 AAAAGATTCTTAAAAATATCAAACTCTGAGATGTTCTGATGTTCCAAATCCAGGAATACA 1120
Db 961 AAAAGATTCTTAAAAATATCAAACTCTGAGATGTTCTGATGTTCCAAATCCAGGAATACA 1020
Qy 1121 GTTGAGGCTCAAGCAGACAGGAAGTAGTATTGGTAGTATTGGTGGTTCCATGCTGTTA 1180
Db 1021 GTTGAGGCTCAAGCAGACAGGAAGTAGTATTGGTAGTATTGGTGGTTCCATGCTGTTA 1080
Qy 1181 GATGATGCTGAAAATGAGACGCTTCCATTTTGTGTTCTGGTGGTTGTCAGATGATTCAC 1240
Db 1081 GATGATGCTGAAAATGAGACGCTTCCATTTTGTGTTCTGGTGGTTGTCAGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATATCTGATTTCTCAAGTGCCCAACAGGAGCTCGCAGCAAGCT 1300
Db 1141 ATGTTCAATACGGAATATCTGATTTCTCAAGTGCCCAACAGGAGCTCGCAGCAAGCT 1200
Qy 1301 AGAGCAGCAAGAGCGCTGAGATGACAGTGTCTGTCAGCGCTGCGAGATGCTCAGAAA 1360
Db 1201 AGAGCAGCAAGAGCGCTGAGATGACAGTGTCTGTCAGCGCTGCGAGATGCTCAGAAA 1260
Qy 1361 GCTTTAGAGCGCTCTAGGTAAAGCTGGGCAACCAAGGAGGATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGCTCTAGGTAAAGCTGGGCAACCAAGGAGGATCTCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTTCTGCTGCTGTGTGAGCGCAGAGTTCCTCCGCTGCGAGCAAGTTCTATA 1480
Db 1321 CAGATCGCTTCTGCTGCTGTGTGAGCGCAGAGTTCCTCCGCTGCGAGCAAGTTCTATA 1380
Qy 1481 GGTCTATCTGTAACAGCTTTACAGACCTCAAAATCTACAGGTTCTGATTTAAACA 1540
Db 1381 GGTCTATCTGTAACAGCTTTACAGACCTCAAAATCTACAGGTTCTGATTTAAACA 1440
Qy 1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTGATGGGCAAGA 1600
Db 1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTGATGGGCAAGA 1500
Qy 1601 AATGATGCACTCGTGTGATGTAACAAATGTAAGTACCCCGCTCTCAGCAGATCCGTT 1660
Db 1501 AATGATGCACTCGTGTGATGTAACAAATGTAAGTACCCCGCTCTCAGCAGATCCGTT 1560
Qy 1661 CTTAGGACAGCAACAGAGCTCGAGGACAGAAAACAGATCAAGCCCTCGCTAGGGTG 1720
Db 1561 CTTAGGACAGCAACAGAGCTCGAGGACAGAAAACAGATCAAGCCCTCGCTAGGGTG 1620
Qy 1721 ATTCTTGGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTCGGCACTCAAACT 1780
Db 1621 ATTCTTGGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTCGGCACTCAAACT 1680
Qy 1781 GTAATGCAATATCCAGTGCATCTTCAAGCAATATGAGGAGATCAGACAAAGCTT 1840
Db 1681 GTAATGCAATATCCAGTGCATCTTCAAGCAATATGAGGAGATCAGACAAAGCTT 1740
Qy 1841 ACATCGGAGTGCAGAAAGCTCCAGCTTTGGCTATCCCTTATGTCACCTTTCTAATGAC 1900
Db 1741 ACATCGGAGTGCAGAAAGCTCCAGCTTTGGCTATCCCTTATGTCACCTTTCTAATGAC 1800

Qy 1901 TCTACACAGAGTTTCATAGCTAAATAGAAAGTTTGTGCTGAAGGATCTAGGACAGCA 1960
Db 1801 TCTACACAGAGTTTCATAGCTAAATAGAAAGTTTGTGCTGAAGGATCTAGGACAGCA 1860
Qy 1961 GCTGAAATAAAAGCAGCTTCTCTTTGAAACCAACTCTTGTGTTTATTCAGCAGGTCGTC 2020
Db 1861 GCTGAAATAAAAGCAGCTTCTCTTTGAAACCAACTCTTGTGTTTATTCAGCAGGTCGTC 1920
Qy 2021 AATATCGGCTCTCTATATTTCTGTTTCTCTCAATAA 2056
Db 1921 AATATCGGCTCTCTATATTTCTGTTTATCTCAATAA 1956

RESULT 6

US-10-312-273-14
; Sequence 14, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035MO
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017989.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 14
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-14

Query Match 90.7%; Score 1956; DB 15; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ATGGTTAATCCTATTGGTCCAGGTCCTATAGCAAGAAACAGAACCCCTCCCGCAGAT 160
Db 1 ATGGTTAATCCTATTGGTCCAGGTCCTATAGCAAGAAACAGAACCCCTCCCGCAGAT 60
Qy 161 CTTTCTGTCTCAAGGATTGGAGCGAGTGCAGCAATAAGAGTGCAGAGCTCAAGAATA 220
Db 61 CTTTCTGTCTCAAGGATTGGAGCGAGTGCAGCAATAAGAGTGCAGAGCTCAAGAATA 120
Qy 221 GCAGTGGGAAGCTAAGCTTAAAGATCTAAGACCGATTCCTAGAGCGATGAGCATC 280
Db 121 GCAGTGGGAAGCTAAGCTTAAAGATCTAAGACCGATTCCTAGAGCGATGAGCATC 180
Qy 281 TTGGTCTTCAGAGTAATGCTCTCATGAGTCTGCAGATAGCTGGGTATTGCTTCTAGT 340
Db 181 TTGGTCTTCAGAGTAATGCTCTCATGAGTCTGCAGATAGCTGGGTATTGCTTCTAGT 240
Qy 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACAGACGACCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACAGACGACCGACCT 300
Qy 401 AGGCTCTCCACCCACCGTTTGTGATTTAAGACTCAGCGCAACAGCTTACGATACT 460

Db 301 |||||ACGCTCTCCACCAGCTTTGATGATTATAGACTCAGCGCAACAGCTTACGATCTACT 360
Qy 461 ATCTTTACCTCAACATCAGTACGTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCT 520
Db 361 ATCTTTACCTCAACATCAGTACGTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCT 420
Qy 521 GTCACTAATATAAGGATACAGCGCTTACTGATCAGGAACCGCAATCGCTGCGGAGTGG 580
Db 421 GTCACTAATATAAGGATACAGCGCTTACTGATCAGGAACCGCAATCGCTGCGGAGTGG 480
Qy 581 GAACTAAGATGCGGATGAGTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 640
Db 481 GAACTAAGATGCGGATGAGTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 540
Qy 641 GCTTCGATACCAAGGATCTGACTCTTTAGTAACTGACTCTTCTCGACTCTTA 700
Db 541 GCTTCGATACCAAGGATCTTGTACTCTTTAGTAACTGACTCTTCTCGACTCTTA 600
Qy 701 CAGGCTGCTCTTCTCCAACTCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAACTCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG 660
Qy 761 ATGCAAGATACCCAGTAGTCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATACCCAGTAGTCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGAT 720
Qy 821 CAGACAGATCTACAGGACACACATAGAGAAAGATGAAATGCTAAATCTAATAACAGTAAAGC 880
Db 721 CAGACAGATCTACAGGACACACATAGAGAAAGATGAAATGCTAAATCTAATAACAGTAAAGC 840
Qy 881 TTTTCAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGC 940
Db 781 TTTTCAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGC 840
Qy 941 AACATAGATTCAGCTAAGCAGCAATGCTACTGCTAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTCAGCTAAGCAGCAATGCTACTGCTAGACACAAATAGCTGAAGCTCAG 900
Qy 1001 AAAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAAATACAGCTGAG 1060
Db 901 AAAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAAATACAGCTGAG 960
Qy 1061 AAAGATCTTAAATAATCAAACTGCGAGATGGTCTGATGTTCCAAATCCAGGAACCTACA 1120
Db 961 AAAGATCTTAAATAATCAAACTGCGAGATGGTCTGATGTTCCAAATCCAGGAACCTACA 1020
Qy 1121 GTTGGAGCTCAAGCAACAGGAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGCTCAAGCAACAGGAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTTA 1080
Qy 1181 GATGATGCTGAAAATGAGACCGCTTCCAAATGCTGTTGATGTTCTGGGTTTCTGATGATTCAC 1240
Db 1081 GATGATGCTGAAAATGAGACCGCTTCCAAATGCTGTTGATGTTCTGGGTTTCTGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATCTGATTTCAAGCTGCCAACAGGAGCTCGCAGCAAGCT 1300
Db 1141 ATGTTCAATACGGAATCTGATTTCAAGCTGCCAACAGGAGCTCGCAGCAAGCT 1200
Qy 1301 AGACGCGGAAGCGCTGGAGATGACAGTCTGCTGCGGCTGGCAGATGCTCAGAAA 1360
Db 1201 AGACGCGGAAGCGCTGGAGATGACAGTCTGCTGCGGCTGGCAGATGCTCAGAAA 1260
Qy 1361 GCTTTAGAGCGGCTCTAGTAAAGCTGGGCAACAAAGGCACTACTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGGCTCTAGTAAAGCTGGGCAACAAAGGCACTACTCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTTCTGCTGCTTTGAGCGCAGAGTTCTCCGCTGCGAGCAAGTTCTATA 1480
Db 1321 CAGATCGCTTCTGCTGCTTTGAGCGCAGAGTTCTCCGCTGCGAGCAAGTTCTATA 1380
Qy 1481 GGGTCATCTGTAACACAGCTTTACAGACCTCAAAATCTACAGGTTCTGATTAATAACA 1540

Db 1381 GGGTCATCTGTAATAACAGCTTTTACAGACCTCAAAATCTACAGGTTCTGATTAATAACA 1440
Qy 1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGATGGGACGA 1600
Db 1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGATGGGACGA 1500
Qy 1601 AATGATGCGACTCGTGTATGATAAAACAATGTAAGTACCCCGCTCTCACACGATCGTT 1660
Db 1501 AATGATGCGACTCGTGTATGATAAAACAATGTAAGTACCCCGCTCTCACACGATCGTT 1560
Qy 1661 CCTAGACGCAACAGAAAGCTCGAGGCCAGAGCAAAACAAGATCAAGCCCTCGCTAGGGTG 1720
Db 1561 CCTAGACGCAACAGAAAGCTCGAGGCCAGAGCAAAACAAGATCAAGCCCTCGCTAGGGTG 1620
Qy 1721 ATTTCGCAATAGCAGAACTCTTGGAGATGCTTAGTCAAGTTTCGGCAGCTACATCT 1780
Db 1621 ATTTCGCAATAGCAGAACTCTTGGAGATGCTTAGTCAAGTTTCGGCAGCTACATCT 1680
Qy 1781 GTAATCAGATCATCCAGTCCAAATCTTCAAGCGAAATAATGAGGAGATCAGACAAAAGCTT 1840
Db 1681 GTAATCAGATCATCCAGTCCAAATCTTCAAGCGAAATAATGAGGAGATCAGACAAAAGCTT 1740
Qy 1841 ACATCGCAGTGACAAAGCTCCACAGTTTGGCTATCTTATGTGCAACTTCTATGAC 1900
Db 1741 ACATCGCAGTGACAAAGCTCCACAGTTTGGCTATCTTATGTGCAACTTCTATGAC 1800
Qy 1901 TCTACACAGAAAGTTCTATAGCTAAATAGAAAGTTTGTTCGTAAGGATCTAGCAGCA 1960
Db 1801 TCTACACAGAAAGTTCTATAGCTAAATAGAAAGTTTGTTCGTAAGGATCTAGCAGCA 1860
Qy 1961 GCTGAAATAAAAGCAGCTTCTTGGAAACGAACTCCCTGTTTATTCAGCAGGTCGTCGTC 2020
Db 1861 GCTGAAATAAAAGCAGCTTCTTGGAAACGAACTCCCTGTTTATTCAGCAGGTCGTCGTC 1920
Qy 2021 AATATCGGCTCTTATATTTCTGCTTATCTCCAAATA 2056
Db 1921 AATATCGGCTCTTATATTTCTGCTTATCTCCAAATA 1956

RESULT 7

US-09-564-479-3
; Sequence 3, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COHEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, JANELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1749)
; US-09-564-479-3

Query Match 85.8%; Score 1848.8; DB 10; Length 1852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 305 ATGAGTCTGCGAGTAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 364

Db 1 ATGAGTCTGGCAGATAAGCTGGTATTCCTTCTAGTAACAGCTGTCTTCTACTAGCAGA 60
QY 365 TCTGACAGAGTGGACTCAACGACAGCGACCTAGCGCTCTCCACCCACGTTTGAT 424
Db 61 TCTGACAGAGTGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCACGTTTGAT 120
QY 425 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTACCTCAACATCACTAGCT 484
Db 121 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTACCTCAACATCACTAGCT 180
QY 485 GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACTAATATAAGGATACAGCG 544
Db 181 GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACTAATATAAGGATACAGCG 240
QY 545 GCTACTGATGAGGAACCGCAATCGCTCGGAGTGGAACTTAAGAAATGCCGATCGATT 604
Db 241 GCTACTGATGAGGAACCGCAATCGCTCGGAGTGGAACTTAAGAAATGCCGATCGATT 300
QY 605 AAAAGTTGGCGCCAAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 664
Db 301 AAAAGTTGGCGCCAAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 360
QY 665 GACTCTTTAGGTAAGTAACTCTTCTTGGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 724
Db 361 GACTCTTTAGGTAAGTAACTCTTCTTGGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 420
QY 725 GCAACAATAACAAAGCAGCTGAGCTCTTAAAGAGATGCAAGATAACCAAGTACTGCCA 784
Db 421 GCAACAATAACAAAGCAGCTGAGCTCTTAAAGAGATGCAAGATAACCAAGTACTGCCA 480
QY 785 GGGAAAAACCCCTGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 844
Db 481 GGGAAAAACCCCTGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
QY 845 ATAGAGAAAGATGGAATGCGATAGGATGATATTTTGGAGCAGAAACCTAGTGA 904
Db 541 ATAGAGAAAGATGGAATGCGATAGGATGATATTTTGGAGCAGAAACCTAGTGA 600
QY 905 GCTGTAGAAAATGCTAAATTAACAGTAAAGCAATAGATTCAGTAAAGCAGCA 964
Db 601 GCTGTAGAAAATGCTAAATTAACAGTAAAGCAATAGATTCAGTAAAGCAGCA 660
QY 965 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCGCCGACTCTCCAAT 1024
Db 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCGCCGACTCTCCAAT 720
QY 1025 CTTCAAGAACCGGAACAAATGTTATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 1084
Db 721 CTTCAAGAACCGGAACAAATGTTATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 780
QY 1085 GCAGATGGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAACAGCA 1144
Db 781 GCAGATGGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAACAGCA 840
QY 1145 AGTAGATTTGGTATGCTGTTTCCATGCTGTTAGATGCTGAAAAATGAGACCGCT 1204
Db 841 AGTAGATTTGGTATGCTGTTTCCATGCTGTTAGATGCTGAAAAATGAGACCGCT 900
QY 1205 TCCATTTTATGCTGCTGGTTTCTGTCAGATGATTCATGTTTCAATACGGAATCTGAT 1264
Db 901 TCCATTTTATGCTGCTGGTTTCTGTCAGATGATTCATGTTTCAATACGGAATCTGAT 960
QY 1265 TCTCAAGCTGCCCAACAGGAGCTCGCAGACCAAGCTTAGACGAGCAAGCCGCTGGAGT 1324
Db 961 TCTCAAGCTGCCCAACAGGAGCTCGCAGACCAAGCTTAGACGAGCAAGCCGCTGGAGT 1020
QY 1325 GACAGTGTCTGCGACGCTGCGAGATGCTCAGAAAGCTTTAGAACGCGCTTAGGTAAA 1384
Db 1021 GACAGTGTCTGCGACGCTGCGAGATGCTCAGAAAGCTTTAGAACGCGCTTAGGTAAA 1080
QY 1385 GCTGGGCAACACAGGCGATCTCAATGTTTAGGACAGATCGCTTCTGCTGTTG 1444

Db 1081 GCTGGGCAACACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
QY 1445 AGCGCAGAGTTCCTCCCGCTGGCAGCAAGTTCCTATAGGGTCATCTGTAAACACGCTTTAC 1504
Db 1141 AGCGCAGAGTTCCTCCCGCTGGCAGCAAGTTCCTATAGGGTCATCTGTAAACACGCTTTAC 1200
QY 1505 AAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGCAGGTTATGATGCT 1564
Db 1201 AAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGCAGGTTATGATGCT 1260
QY 1565 TACAATTCATCAATGATGCTATGTTAGGCGCAAGAAATGATGCGACTCGTGTGATGATA 1624
Db 1261 TACAATTCATCAATGATGCTATGTTAGGCGCAAGAAATGATGCGACTCGTGTGATGATA 1320
QY 1625 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGACACGAAACAGAAAGCTCGA 1684
Db 1321 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGACACGAAACAGAAAGCTCGA 1380
QY 1685 GGACCAAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGCAATAGCAGAACTCTT 1744
Db 1381 GGACCAAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGCAATAGCAGAACTCTT 1440
QY 1745 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAAATGCGAGATCATCCAGTCGAT 1804
Db 1441 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAAATGCGAGATCATCCAGTCGAT 1500
QY 1805 CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGCACTGACAAAAGCTCCA 1864
Db 1501 CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGCACTGACAAAAGCTCCA 1560
QY 1865 CAGTTTGGCTATCTTATGTGCAACTTTCTAATGACTCTACACAGAAAGTTCTATAGCTAAA 1924
Db 1561 CAGTTTGGCTATCTTATGTGCAACTTTCTAATGACTCTACACAGAAAGTTCTATAGCTAAA 1620
QY 1925 TTAGAAAGTTTGTGCTGAGGATCTAGGACAGCTGAAATGAAAGCACTTTCCTTT 1984
Db 1621 TTAGAAAGTTTGTGCTGAGGATCTAGGACAGCTGAAATGAAAGCACTTTCCTTT 1680
QY 1985 GAAACGAACTCCTTGTATTATTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATTCTGGT 2044
Db 1681 GAAACGAACTCCTTGTATTATTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATTCTGGT 1740
QY 2045 TATCTCCAATAACAAACACTTAAGTGTGCTTTGGAGAGATTTATGTTGCTTTGTTAAGG 2104
Db 1741 TATCTCCAATAACAAACACTTAAGTGTGCTTTGGAGAGATTTATGTTGCTTTGTTAAGG 1800
QY 2105 CTTTGTGTGAGGCTTACCAACACACTAGACGATCTTCAATAAATAAAGA 2156
Db 1801 CTTTGTGTGAGGCTTACCAACACACTAGACGATCTTCAATAAATAAAGA 1852

RESULT B

US-09-564-479-7
; Sequence 7, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COHEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: US/09/564,479
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR FILING DATE: 1999-05-03
; PRIOR FILING DATE: 1999-06-30
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2238

```
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (766)..(2235)
US-09-564-479-7

Query Match      68.0%; Score 1465.8; DB 10; Length 2238;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1478; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTAGAGAGCTTTTATTTTATTTT 60
Db 666 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTAGAGAGCTTTTATTTTATTTT 725

QY 61 ATATATAACTATAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db 726 ATATATAACTATAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 785

QY 121 AGGTCTTATAGACGAAACAGAACGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA 180
Db 786 AGGTCTTATAGACGAAACAGAACGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA 845

QY 181 GCGAGTGCAGCAATAATAGAGTGGGAAGCTCAAGAATAGCAGGTGCGGAAGCTTAAGCC 240
Db 846 GCGAGTGCAGCAATAATAGAGTGGGAAGCTCAAGAATAGCAGGTGCGGAAGCTTAAGCC 905

QY 241 TAAAGATCTAAGACCGGATCTGTAGCGATGGAGCATCTTGGTCTCGAGTGAATGC 300
Db 906 TAAAGATCTAAGACCGGATCTGTAGCGATGGAGCATCTTGGTCTCGAGTGAATGC 965

QY 301 TCTCATAGTCTGCGAGATAAGCTGGGTATTGCTTTAGTAAACAGCTCGTCTTCTACTAG 360
Db 966 TCTCATAGTCTGCGAGATAAGCTGGGTATTGCTTTAGTAAACAGCTCGTCTTCTACTAG 1025

QY 361 CAGATCTCAGACGTGGTGAATCAAGACAGACGACCGCCTAGCCTCTCTCACCCAGCTT 420
Db 1026 CAGATCTCAGACGTGGTGAATCAAGACAGACGACCGCCTAGCCTCTCTCACCCAGCTT 1085

QY 421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACT 480
Db 1086 TGATGATTATAGACTCAAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACT 1145

QY 481 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCTCAGGATGCTGTCACTAATAAAGGATAC 540
Db 1146 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCTCAGGATGCTGTCACTAATAAAGGATAC 1205

QY 541 AGCGCTACTGATGAGGAAACCGGATCGCTGCGAGTGGGAATAGAAATGCGGATGC 600
Db 1206 AGCGCTACTGATGAGGAAACCGGATCGCTGCGAGTGGGAATAGAAATGCGGATGC 1265

QY 601 AGTTAAAGTTGCGGCGCAATTTACAGATTAGCGAAATATGCTTCGGATAACCAAGCGAT 660
Db 1266 AGTTAAAGTTGCGGCGCAATTTACAGATTAGCGAAATATGCTTCGGATAACCAAGCGAT 1325

QY 661 TCTTGACTCTTTAGTAAATGACTTCTTCACTCTTCAAGGCTGCTCTTCTTCCAAATC 720
Db 1326 TCTTGACTCTTTAGTAAATGACTTCTTCACTCTTCAAGGCTGCTCTTCTTCCAAATC 1385

QY 721 TGTAGCAACATATAACAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db 1386 TGTAGCAACATATAACAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 1445

QY 781 CCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC 840
Db 1446 CCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC 1505

QY 841 ACAGATAGAGAAGATGGAATTCGATTTAGGATGCATTTTTCAGGACAGAACGCTAG 900
Db 1506 ACAGATAGAGAAGATGGAATTCGATTTAGGATGCATTTTTCAGGACAGAACGCTAG 1565

QY 901 TGGAGCTCTAGAAATGCTTAATCTAATAACAGTATAGCAACATAGATTCACTTAAGCC 960
Db 1566 TGGAGCTCTAGAAATGCTTAATCTAATAACAGTATAGCAACATAGATTCACTTAAGCC 1625

QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGGACTCTCC 1020
Db 1626 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGGACTCTCC 1685

QY 1021 AATTCTTCAAGAAGCGGAACAAATGGTAATACAGGCTGGAAGAGATCTTAAAAATATCAA 1080
Db 1686 AATTCTTCAAGAAGCGGAACAAATGGTAATACAGGCTGGAAGAGATCTTAAAAATATCAA 1745

QY 1081 ACCTGCAGATGCTTCTGATGTTCCAAATCCAGNACTACAGTTGGAGGCTCCAGGCAACA 1140
Db 1746 ACCTGCAGATGCTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGGCAACA 1805

QY 1141 AGGAAGTAGTATTGGTAGTATTGCTGTTCCATGCTCTAGATGATGCTGAAAAATGAGAC 1200
Db 1806 AGGAAGTAGTATTGGTAGTATTGCTGTTCCATGCTCTAGATGATGCTGAAAAATGAGAC 1865

QY 1201 CGCTTCCATTTTGTATGCTGGGTTTCGTGATGATTCATGTTCAATGCTTGAATGAGAC 1260
Db 1866 CGCTTCCATTTTGTATGCTGGGTTTCGTGATGATTCATGTTCAATGCTTGAATGAGAC 1925

QY 1261 TGATTTCAAGCTGCTCCCAACAGGAGCTCGCAGCAACAGCTAGACGCGAAGCGCTGG 1320
Db 1926 TGATTTCAAGCTGCTCCCAACAGGAGCTCGCAGCAACAGCTAGACGCGAAGCGCTGG 1985

QY 1321 AGATGACAGTGTGCTGCTGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGCGCTTAGG 1380
Db 2045 AGATGACAGTGTGCTGCTGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGCGCTTAGG 2045

QY 1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Db 2046 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 2105

QY 1441 TGTGAGCGCAGGAGTTCCTCCGCTGCGAGCAAGTTCTATAG 1481
Db 2106 TGTGAGCGCAGGAG-TACTCCGCTGCGAGCAAGTTCTATAG 2145

RESULT 9
US-09-564-479-5
; Sequence 5, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OOMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1456
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1456)
US-09-564-479-5

Query Match      67.5%; Score 1455; DB 10; Length 1456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTAGAGAGCTTTTATTTTATTTT 60
```


Db 483 CGCAGGAACAGCTGCGCACTGTCCAGATGAATGTATAAACAGCTTTTACAGACAGCGTTT 542
Qy 1518 CTACAGGTTCT-----GATTATAAACACAGATATCACAGGTTATGATGCTTACAAT 1571
Db 543 CTTGCACTTCTTCCAGCTCTTATGACGAGCACTTTCCGATGGATNTCTGTTACAAA 602
Qy 1572 CCATCAATGATCGCTATGTTAGGCGCACGAAATGATCGCACTCGTGATGTGATAAACAAATG 1631
Db 603 CACTGAACCTTTATATTCGGAACAGAGCGCGTGCACTATATAGTCAACCTG 662
Qy 1632 TAAATACCCCGCTCTCAACAGTCCGTTCTTAGAGCAGCAAGAGCTCCAGACCAAG 1691
Db 663 CAAATCCCGCTTTCCAGAGCGTTTCTCGTTCTGGCATAGAAAGCTCAAGGACGCAATG 722
Qy 1692 AAAAAACAGATCAAGCCCTCGTAGGTTGATTTCTGCAATAGCAAACTCTTGGAGATG 1751
Db 723 CAGATGCTAGCCAAAGAGCAGCAAGAACTATTGTGAGATAGCCAAAGCTTAGGTGATG 782
Qy 1752 TCTATAGTCAAGTTTCCGCACTACAATCTGTAATGAGATCATCCAGTCAATCTCTCAAG 1811
Db 783 TATATAGCCGTTTACAGGTTCTGGATTCTTGTATGCTACGATTGTGAGCAATCCGCAAG 842
Qy 1812 CGAATATAGAGATCAGCAAAAGCTTACATCGGCACTGACAAAGCTCCACAGTTTG 1871
Db 843 CAAATCAAGAAGAGATTATGCAAGAGCTCACGGCATCTATTAGCAAGCTCCACAATTTG 902
Qy 1872 GCTATCCTTATGTGCAACTTCTTAATGACTCTACACAGAGTTTCATAGCTTAATAGAA 1931
Db 903 GGTATCCTCTGTTAGAAATCTGTTGGATAGCTTGCAAGAGTTTGTCGCAATTTGAAA 962
Qy 1932 GTTGTCTCTGAAGATCTAGACAGAGCTGAAATAAAGCACTTTCCTTTGAAACGA 1991
Db 963 GAGAGTTTGTGATGGGAAAGCTAGTCTCGCAATCTCAAGAGAACTGTTTAGAAAC 1022
Qy 1992 ACTCTTGTGTTATTCAGAGGTTCTGCTCAATCGGCTCTCTATATCTGTTATCTCC 2051
Db 1023 AGCCCGCTTCTATCAACAGGTTGGTAAACATGCTTCTATCTCTGTTATCTTT 1082
Qy 2052 AATAAC 2057
Db 1083 CTTAAC 1088

RESULT 11

US-09-841-132-412

; Sequence 412, Application US/09841132

; Patent No. US20020061848A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841,132

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 412

; LENGTH: 1941

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis serovar D

US-09-841-132-412

Query Match 8.6%; Score 185.4; DB 9; Length 1941;
Best Local Similarity 48.5%; Pred. No. 5e-38;

Matches 717; Conservative 0; Mismatches 721; Indels 39; Gaps 6;

Qy 606 AAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATCTCTCGGATAACCAAGCGATTCTTG 665

Db 467 AGGTATTGATAGCTCAATCATAGTTAGTTAACTTGAAAACAGAAATCAGACTTTAAAGG 526

Qy 666 ACTCTTTAGGTAAACCTGACTTCCTTCGACCTCTTTACAGGCTGCTCTTCTCCAATCTGTAG 725
Db 527 AAACCTTTTAAACACACAGACTCTGCAGATCAGATTCACAGGATTAATAGTCAGTTAGAGA 586
Qy 726 CAAACAATTAACAAGCAGCTGAGCTTCTTAAAGAGATGAAGATAACCCAGTAGTCCACAG 785
Db 587 TCAACAAAAATTTCTGCAGATCAAAATTTATCAAAAGATCTGGAAGGACAAAACATAAGTTATG 646
Qy 786 GGAACCCCTCGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACACAGA 845
Db 647 AAGCTGTTCTCACTAACCCAGAGAGGTTATCAAAAGCTTCTTCTGAAGCGGAATTAAGT 706
Qy 846 TAGAGAAAAGATGGAATTCGATTTAGGATGATATTTTGCAGGACAGAACGCTAGTGGAG 905
Db 707 TAGGACAAGCTTTGCGAGTCTATTTGTGGATGCTGGGGATCAAAAGCCAGGCTCGAGTTCTTC 766
Qy 906 CTGTAGAAAATATGCTAAATCTTAATTAACGATTAAGCAACATAGATTCAGCTAAGCAGAA 965
Db 767 AAGCAGCAAAATTAATAGCCAGATAATATCGCAGCCACGAAGAAATTAATTTGATGCTG 826
Qy 966 TCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCCAATTC 1025
Db 827 CTGAAAACGAGGTAACGAGTTAAACGAGCATACAGGGCTAACGGACTCGCCTTTAG 886
Qy 1026 TTCAAGAAAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTTAAAAATATCAAACTG 1085
Db 887 TGAATAAAGCTGAGGAGCAGATTAGTCAAGCACAATAAGATATTCAGAGATCAAACTT- 945
Qy 1086 CAGATGTTCTGATGTTCCAAATCCAGGAATCAAGGAACTACAGTTGGAG-----GCTCAAGCAACA 1140
Db 946 --AGTGGTTCGGATATTCCTATCGTTGTCGAGTGGGTGAGTCTTCCGAGGAAGTG 1003
Qy 1141 AGGAAGTAGTATTTGGTAGTAT-----TCGTGTTTCCATGCTGTTTGTAGATGATG 1187
Db 1004 CGGTAGGAGCGTTGAAATCTCTAACTTCAAGAGAAATTTCTTGTGTTGATGATG 1063
Qy 1188 CTGAAAATGAGACCGCTTCCATTTTGTATGTTGCGGTTTCGTGAGATGATTCATGTTTCA 1247
Db 1064 TAGACAATGAAATGGCAGCGATTGCAATGCAAGGTTTTCGATCTATGATCGAACAAATTA 1123
Qy 1248 ATACGGAATTCCTGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGAGCAG 1307
Db 1124 ATGTAAACAATCTCTGCAACAGCTAAAGAGCTTACAGCTATGGAGCTCAGCTGA---CTG 1180
Qy 1308 CGAAAGCCGCTGGAGATGACAGATGCTGCTGAGCGCTCGCAGATGCTCAGAAAGCTTTAG 1367
Db 1181 CGATGTCAGATCAACTGTTGTTGGTGGGATGGGAGCTCCGAGCCGAAATACAGCAATCA 1240
Qy 1368 AAGCGCTCTAGGTAAGCTGGGCAACAAAC-----AGGGCATACTCAATGCTTTAG 1418
Db 1241 AAGATGCTCTTGGCAAGCTTTGAAAACCAACCATCAACAGATGGTTTAGTACAGCTATGG 1300
Qy 1419 GACAGATCGCTCTGCTGCTGTTGTGAGCGCAGGAGTTCTTCCGCTCAGCAAGTTCTTA 1478
Db 1301 GACAAAGTGGCTTTTGCACTGCTCCAGGTTGGAGGAGGCTCCGAGGAACAGCTGGCACTG 1360
Qy 1479 TAGGTCATCTGTAAAAACAGCTTTTACAAGACTCAAAAATCTACAGGTTCTGA-----TT 1532
Db 1361 TCCAGATGAATGTAATAACAGCTTTTACAAGACAGCTTTTCTTCGACTTCTTCCAGCTCTT 1420
Qy 1533 ATAAACACAGATATCAGAGGTTATGATGCTTTTCAAAATCCATCAATGATGCTTATGTA 1592
Db 1421 ATGACAGCAGCACTTTTCCGATGATTTCTGCTTACAAAACACTGAACTCTTTATATTCG 1480
Qy 1593 GGGCAGGAATGATGCGACTGCTGATGTAATAACATGTAAGTACCCCGCTCTCACAC 1652
Db 1481 AAAGCAGAAAGCGGCTGCGATGATGTAAGTCAAGAGCGAGTGCAGATGCTAGCCAAAGAGCAG 1540
Qy 1653 GATCCGTTCTCTAGACAGCAAGAGCTCGAGGACCAAGAAAAACAGATCAAGCCCTCG 1712
Db 1541 CGGTTTCTCGTTCTGGCATAGAAAGTCAAGAGCGAGTGCAGATGCTAGCCAAAGAGCAG 1600
Qy 1713 CTAGGGTGATTTCTTGGCAATAGCAGAACTCTTGGAGATGCTCTATAGTCAAGTTTTCGGCAC 1772

Db 1601 CAGAACTATTGTGACAGATAGCCAAAGCTTAGGTGATGATATAGCCGCTTACAGGTTT 1660
QY 1773 TACAATCTGTAATCAGATCAATCCAGTCCGAAATCCTCAAGCGCAATTAATGAGGATCAGAC 1832
Db 1661 TGGATTCTTTGATGCTACGATTTGTGAGCAATCCGCAAGTAAATCAAGAAGAGATTATGC 1720
QY 1833 AAAAGCTTACATCGGAGTGCAGAGCTTCCACAGCTTTGGCTATCCTTATGTCGAATTT 1892
Db 1721 AGAAGCTCAGGCACTCTATTAGCAAGCTCCCAATTTGGGTATCCTGCTGTTCAGAATT 1780
QY 1893 CTAATGACTCTACACAGAGTTTCATAGCTTAAATTAGAAAAGTTTGTGCTGAAGGATCTA 1952
Db 1781 CTGCGGATAGCTTCAGAGATTTCGCGCAATTGGAAGAGAGTTTGTGATGGGAAC 1840
QY 1953 GGACAGAGCTGAATATAAAGCACTTTCTTTGAAACGAACTTCCTTTTATTCAGCAGG 2012
Db 1841 GTAGTCTCGCAGATCTTCGAGAGAATGCGTTTGTAGAAAACAGCCCGCTTTTCATTCACAGG 1900
QY 2013 TGCTGGTCAATATCGCTCTCTATATCTCGTTATCT 2049
Db 1901 TGTGGTAAACATGCTTCTCTATCTCTCGGTTATCT 1937

RESULT 12

US-09-841-260-13

; Sequence 13, Application US/09841260

; Publication No. US20030175700A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Probst, Peter

; APPLICANT: Stromberg, Erika Jean

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS

; FILE REFERENCE: 210121.515

; CURRENT APPLICATION NUMBER: US/09/841,260

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 140

; SEQ ID NO 13

; LENGTH: 1171

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-841-260-13

Query Match 8.5%; Score 184; DB 10; Length 1171;
Best Local Similarity 52.8%; Pred. No. 8.2e-38;
Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;
QY 1167 TTCCATGCTGTAGATGATGCTGAAATGAGACCGCTTCCATTTTGTGATGCTGGGTTTC 1226
Db 207 TTTCTTGTGCTGATGATGATGACATGAAATGGCAGCGATTGCACTGCAAGGTTTTC 266
QY 1227 GTCAGATGATTCACATGTTCAATACGGAATCCCTGATTTCTACAGTCCGCAACAGGAGC 1286
Db 267 GATCTATGATCGCAACAATTTAATGTAAACATCTCGCAACAGCTAAAGAGCTACAGCTA 326
QY 1287 TCGCAGCAAGCTAGACAGCGCAAGCGCTGGAGATGACAGTGTGCTGAGCGCTGG 1346
Db 327 TGGAGCTCAGCTGA---CTGCAATGTCAGATCACTGTTGTTGGGATGGCAGCTCC 383
QY 1347 CAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGGTAAAGCTGGGCAACAAAC----- 1397
Db 384 CAGCCGAAATACAAAGCAATCAAGATGCTCTTGGCGAAGCTTTGAAACAAACCATCAGCAG 443
QY 1398 AGGGCACTACATGCTTTTAGGACAGATCGCTTCTGCTGCTGTTGTGAGCGAGGATTC 1457
Db 444 ATGGTTTGGCTACAGCTATGGGCAAGTGGCTTTTTCAGCTGCCAAGGTTGAGGAGGCT 503
QY 1458 CTCGCCCTGCAGCAAGCTTCTATAGGCTCATCTGTAAACAGCTTTTCAAGACCTCAAAAT 1517
Db 504 CCGCAGAACAGCTGCACTGCTCCAGATGATGATGATGATGATGATGATGATGATGATGAT 563
QY 1518 CTACAGGTTCT-----GATTATAAAACAGATATCAGCAAGGTTATGATGATGATGAT 1571

Db 564 CTTGCACTTCTTCCAGCTCTTATGACAGACACTTTCCGATGGATATCTGCTTACAAA 623
QY 1572 CCATCAATGATGCTATGGTAGGSCACGAATGATGAGCTCGGTGATGATGATGATGATGAT 1631
Db 624 CACTGAACCTCTTTATATTCCGAAAGCAGAAAGCGGCTGCACTAGCTATTAGTCAAACTG 683
QY 1632 TAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAGCTCGAGGACGAC 1691
Db 684 CAAATCCCGCTTCCAGAAAGCTTCTTCTTCTGCGATAGAAAGTCAAGAGCGAGT 743
QY 1692 AAAAAACAGATCAAGCCCTCGTAGGTGATTTCTTGGCAATAGCAGAACTCTTGGAGATG 1751
Db 744 CAGATGCTAGCCAAAGAGCAGCAGAAACTATTGTGACAGATAGCCAAAGCTTAGGTGATG 803
QY 1752 TCTATAGTCAAGTTTGGCACTACATCTGTAATGAGATCATCCAGTCCGATCTCAAG 1811
Db 804 TATATAGCGGCTTACAGGTTCTGGAATCTTTGATGTTCTACGATTTGTGAGCAATCCGCAAG 863
QY 1812 CGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTCAAAAGCTCCACAGTTTG 1871
Db 864 CAAATCAGAGAGATTTATGCAAGAGCTCAGCGCATCTATTAGCAAGCTCCCAATTTG 923
QY 1872 GCTATCTTATGTCAACTTTCTAATGACTCTACACAGAGTTTCAATAGTAAATAGAAA 1931
Db 924 GGTATCTCTGCTGTTTCAAGATTTCTGGGATAGCTTGCAGAACTTTGCTGCGCAATTTGAAA 983
QY 1932 GTTTGTTTGTGAGGATCTAGGACAGCAGCTGAAATAAAGCAGCTTCTCTTGAACGA 1991
Db 984 GAGAGTTTGTGATGGGGAAGCTAGTCTCGCAGAACTCTCAAGAGAACTCGGTTTGAACAA 1043
QY 1992 ACTCTCTTATTTATTCAGCAGGCTGCTGCTCAATATCGGCTCTCTATATTCTGTTATCTCC 2051
Db 1044 AGCCGCTTCTTCAACACAGGTTTGGTAAATGCTTCTCTATCTCTGTTATCTTT 1103
QY 2052 AATAAC 2057
Db 1104 CTTAAC 1109

RESULT 13

US-10-007-693-13

; Sequence 13, Application US/10007693

; Publication No. US20020146776A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT

; FILE REFERENCE: 210121.515C2

; CURRENT APPLICATION NUMBER: US/10/007,693

; CURRENT FILING DATE: 2001-12-05

; NUMBER OF SEQ ID NOS: 157

; SEQ ID NO 13

; LENGTH: 1171

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-10-007-693-13

Query Match 8.5%; Score 184; DB 13; Length 1171;
Best Local Similarity 52.8%; Pred. No. 8.2e-38;
Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;
QY 1167 TTTCCATGCTGTAGATGATGCTGAAATGAGACCGCTTCCATTTTGTGATGCTGGGTTTC 1226
Db 207 TTTCTTGTGCTGATGATGATGACATGAAATGGCAGCGATTGCACTGCAAGGTTTTC 266
QY 1227 GTCAGATGATTCACATGTTCAATACGGAATCCCTGATTTCTCAAGCTCCCAACAGGAGC 1286
Db 267 GATCTATGATCGCAACAATTTAATGTAAACATCTCGCAACAGCTAAAGAGCTACAGCTA 326
QY 1287 TCGCAGCAAGCTAGACAGCGCAAGCGCTGGAGATGACAGTGTGCTGAGCGCTGG 1346

Db 327 TGGAGGCTCAGCTGA---CTGGATGTCAGATCAACTGGTTGGTGGGATGGGAGCTCC 1383
QY 1347 CAGATGCTCAGAAAGCTTTAGAGCGGCTTAGTAAAGCTGGGCAACAAAC----- 1397
Db 384 CAGCGAAATCAAGCAATCAAGATGCTCTTGGCGAAGCTTTGGAACAAACATCAGCAG 443
QY 1398 AGGGCATACTCAATGCTTTAGGACATCGCTTCTGCTCTGTTGGAGCGCAGAGTTTC 1457
Db 444 ATGGTTGGCTACAGCTATGGACAGTGGCTTTTGACGCTGCCAGGTTGGAGAGGCT 503
QY 1458 CTCGCGCTCCAGCAAGTTCTATAGGTCATCTGTAAACAGCTTTACAGACCTCAAAAT 1517
Db 504 CGCGAGGAACAGCTGCACTGTCCAGATCAATGTATGAACAGCTTTACAGACAGCGTTT 563
QY 1518 CTACAGGTTCT-----GATTATAAACACAGATATCAGCAGGTTATGATGTTCAAAAT 1571
Db 564 CTTGACATCTTCCAGCTCTTATGACAGCAGCTTTCCGATGGATATCTGTTTACAAA 623
QY 1572 COATCAATGATGCCTATGGTAGGCGACGAAATGATGCGACTCGTGTATGTATAAAACAATG 1631
Db 624 CACTGAATCTTTATATTCGAAAGCAGAAAGCGGCTGCAGTCAGCTATTAGTCAAAC 683
QY 1632 TAAGTACCCCGCTCTCACACGATCGGTTCTTAGAGCAGCAACAGAGCTCGAGGACGAG 1691
Db 684 CAAATCCCGCGCTTTCCAGAAAGCTTTCTCGTTCTGGCATAGAAAGTCAAGGACGAGTG 743
QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATAGCAGAACTCTTGAGAGTG 1751
Db 744 CAGATGCTAGCCAAAGACGACGAGAAATATTTGTGAGATGACCAAGCTTAGTGATG 803
QY 1752 TCTATAGTCAAGTTTGGCCTACAACTGTAATGATGATGATGATGATGATGATGATGAT 1811
Db 804 TATATAGCGCTTACAGGTTCTGGATTTCTTGTGCTAGCTGATGATGATGATGATGATG 863
QY 1812 CGAATAAGTACAGATCAGCAAAAGCTTACATGGCAGTGAACAGCTTCCAGAGTTTG 1871
Db 864 CAAATCAAGAGAGATATGACAGAGCTCACGGCATCTATTAGCAAGCTCCACAAATTTG 923
QY 1872 GCTATCCTTATGCTCAACTTTCTAATGACTTACACAGAGTTTCATAGTAAATAGAAA 1931
Db 924 GGTATCCTGCTGTTAGAAATCTGGGATAGCTTGCAGAGTTTGGTGGCAATTTGAAA 983
QY 1932 GTTGTGCTGAAGATCTAGGACAGCAGCTGAAATAAAGCACTTTCTTTGAAACGA 1991
Db 984 GAGAGTTTGTGATGGGAAACGTAGTCTCGCAGAACTTCAAGAGAAATGGTTTGAAC 1043
QY 1992 ACTCTGTTTATCAGAGGCTGCTGCTCAATATCGGCTCTATATCTGTTTATCTCC 2051
Db 1044 AGCCGCTTTCATCAACAGAGTTTGGTAAACAATGCTTCTATCTCTGTTATCTTT 1103
QY 2052 AATAAC 2057
Db 1104 CTTAAC 1109

RESULT 14
US-09-841-260-38
; Sequence 38, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probet, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 38
; LENGTH: 1834
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis

US-09-841-260-38

Query Match 8.5%; Score 184; DB 10; Length 1834;
Best Local Similarity 52.8%; Pred. No. 1.1e-37;
Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;

QY 1167 TTTCCATGCTGTAGATGATGCTGAAATGAGACCGCTTCCATTTTGTATGTCGGTTTC 1226
Db 816 TTTCCCTGTTGCTTGATGATGTAGCAATGAAATGGCAGCGATTCATCTGCAGGTTTC 875
QY 1227 GTCAGATGATTCAATGTTCAATACGGAATTCCTGATTTCTCAAGCTCCCAACAGGAGC 1286
Db 876 GATCTATGATCGAAATTTTAAATGTAAACATCTCTGCAACAGCTAAAAGCTACAGCTA 935
QY 1287 TCGCAGCAAGAGCTTAGACGAGGAAAGCGGTGGAGATGACAGTCTGCTGAGCGCTGG 1346
Db 936 TGGAGGCTCAGCTGA---CTGCGATGTGAGATCAACTCGTTGGTGGGATGGCAGCTCC 992
QY 1347 CAGATGCTCAGAAAGCTTTAGAAGCGGCTCTAGGTAAAGCTGGGCAACAAC----- 1397
Db 993 CAGCGAAATACAGCAATCAAGATGCTCTTGCCAAAGCTTTGAAACAACCTCAGCAG 1052
QY 1398 AGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGATTC 1457
Db 1053 ATGGTTTGGCTACAGCTATGGGCAAGTGGCTTTTGCAGCTGCCAAGGTTGGAGGAGCT 1112
QY 1458 CTCGCGCTGACGCAAGTTCTATAGGTCATCTGTAAACAGCTTTTACAGACCTCAAAAT 1517
Db 1113 CCGCAGGAAACAGCTGGGCTCTCCAGATGAATGTAAACAGCTTTTACAGACAGCGTTT 1172
QY 1518 CTACAGGTTCT-----GATTATAAACACAGATATCAGAGGTTATGATGTTTACAAT 1571
Db 1173 CTTGACATCTTCCAGCTCTTATGCGACGACACTTTCGATGGATATCTCTGTTACAAA 1232
QY 1572 CCAATCAATGATGCTGATGGGCAAGAAATGATGCGACTCGTGTATGATAACAATG 1631
Db 1233 CACTGAATCTTTATATTCGAAAGCAGAGCGGCTGCAGTCACTATTAGTCAACCTG 1292
QY 1632 TAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGACAGAGCTCGAGGACGAG 1691
Db 1293 CAAATCCCGCGCTTCCAGAGCGCTTTCTCGTTCTGGCATAGAAAGTCAAGGACGAGTG 1352
QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGCAATAGCAGAACTCTTGGAGATG 1751
Db 1353 CAGATGCTAGCCAAAGACAGCAAGAACTATTGTGAGATAGCCAAACGTTAGGTGATG 1412
QY 1752 TCTATAGTCAAGTTTGGCACTA CAATCTGTAATGACAGATCATCCAGTCGAATCTCTAAG 1811
Db 1413 TATATAGCGGCTTACAGGTTCTGGATTTCTTTGATGTTTACGATTGTGAGCAATCCGCAAG 1472
QY 1812 CGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCTCCACAGTTTG 1871
Db 1473 CAAATCAGAAAGATATGACAGAGCTCAGCGCATCTATTAGCAAGCTCCACAAATTTG 1532
QY 1872 GCTATCCTTATGTCAACTTTCTAATGACTCTACACAGAACTTCATAGCTAAATAGAAA 1931
Db 1533 GGTATCCTGCTGTTGAGAAATTTCTCGGATAGCTTTGAGAAAGTTGCTGGCAATTTGAAA 1592
QY 1932 GTTGTGTTGTAAGGATCTAGGACAGAGCTGAAATAAAGCACTTTCTTTGAAACGA 1991
Db 1593 GAGAGTTTGTGATGGGAAACGTAGTCTCGCAGAACTCTCAAGAGATGCGTTTAAAGAAC 1652
QY 1992 ACTCCTTGTATTACAGAGCTGCTGCTCAATATCGGCTCTCTATATCTGTTTATCTCC 2051
Db 1653 AGCCGCTTTCATCAACAGAGTTTGGTAAACAATGCTTCTTCTTCTGTTTATCTTT 1712
QY 2052 AATAAC 2057
Db 1713 CTTAAC 1718

RESULT 15
US-10-007-693-38

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:23:38 ; Search time 160.201 Seconds
(without alignments)
7067.770 Million cell updates/sec

Title: US-10-608-559-1

Perfect score: 2156

Sequence: 1 ataaatctttaaaacagg.....gatctcaataaaaaaga 2156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1155024 seqs, 262583689 residues

Total number of hits satisfying chosen parameters: 2310048

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_New.*

1: /cgn2_6/prodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/prodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/prodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/prodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/prodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/prodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	49	2.3	1860	6	US-10-414-532-71	Sequence 71, Appl
2	49	2.3	2000	6	US-10-414-532-20	Sequence 20, Appl
3	45.8	2.1	775	6	US-10-414-532-29	Sequence 29, Appl
4	45.8	2.1	1473	6	US-10-414-532-33	Sequence 33, Appl
5	41.2	1.9	416	6	US-10-767-701-16328	Sequence 16328, A
6	41	1.9	234309	6	US-10-796-280-12378	Sequence 12378, A
7	40	1.9	1790242	6	US-10-767-471-10805	Sequence 10805, A
8	39.6	1.8	447	6	US-10-767-701-17775	Sequence 17775, A
9	39.6	1.8	466	6	US-10-767-795-178	Sequence 178, App
10	39.2	1.8	165223	6	US-10-796-280-12544	Sequence 12544, A
11	38.8	1.8	335	6	US-10-793-479-31955	Sequence 31955, A
12	38.8	1.8	318752	6	US-10-767-471-10553	Sequence 10553, A
13	38.4	1.8	52746	6	US-10-765-790-39	Sequence 39, Appl
14	38.4	1.8	119006	7	US-60-550-051-3011	Sequence 3011, Ap
15	38.2	1.8	2685	6	US-10-784-480-357	Sequence 357, App
16	38	1.8	770	1	PCT-US04-05757-19	Sequence 19, Appl
17	38	1.8	770	1	PCT-US04-05757-19	Sequence 19, Appl
18	38	1.8	967	1	PCT-US04-05757-18	Sequence 18, Appl
19	38	1.8	967	1	PCT-US04-05757-18	Sequence 18, Appl
20	38	1.8	1160	1	PCT-US04-05757-17	Sequence 17, Appl
21	38	1.8	1160	1	PCT-US04-05757-17	Sequence 17, Appl
22	38	1.8	1367	1	PCT-US04-05757-16	Sequence 16, Appl
23	38	1.8	1367	1	PCT-US04-05757-16	Sequence 16, Appl
24	38	1.8	1553	1	PCT-US04-05757-15	Sequence 15, Appl
25	38	1.8	1553	1	PCT-US04-05757-15	Sequence 15, Appl
26	38	1.8	1719	1	PCT-US04-05757-14	Sequence 14, Appl

RESULT 1
US-10-414-532-71
; Sequence 71, Application US/10414532
; GENERAL INFORMATION:
; APPLICANT: CURTISS III, ROY
; APPLICANT: KANG, HO YOUNG
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING
; FILE REFERENCE: 56029-40437
; CURRENT APPLICATION NUMBER: US/10/414,532
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/372,710
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 71
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
US-10-414-532-71

Query Match 2.3%; Score 49; DB 6; Length 1860;
Best Local Similarity 49.5%; Pred. No. 0.0073;
Matches 155; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

QY	845	ATAGAGAAAGATGGAAATGCGATAGGATGCATATTTTTCAGGACAGACGCTAGTGGGA	904	Sequence 14, Appl
Db	346	ACAGACAAAGCGGCAAAAGACGACAGATAGATGATAGATGAAGCTAAGAAACGCGAA	405	Sequence 13, Appl
QY	905	GCTGTAGAAATGCTTAATCTAATAACAGTATAGCAACATAGATTAGCTTAAGCAGCA	964	Sequence 13, Appl
Db	406	GAAAGGCAAAACATAATTTAATCTGTTTCGACCAATGGTAGTTCCTGAGCCAGCAG	465	Sequence 1, Appl
QY	965	ATCGTCTACTGTAAAGACACAATAGCTGAAGCTCAGAAAAAGTTCCCGGA---CTCTCCA	1021	Sequence 1, Appl
Db	466	TTGCTGAGACTAAGAAAAATCAGAGAGAGCTAATAAAGACACAGAACTTACTAAA	525	Sequence 77, Appl
QY	1022	ATTCCTCAAGAGCGGAAACAAATGGTAATACAGCTGAGAAAGATCTTAAATAATCAAA	1081	Sequence 19408, A
Db	526	AAACTAGAGAGAGCTAAAGCAAAAATTAGAGAGGCTGAGAAAAAGCTACTGAAGCCAAA	585	Sequence 23459, A
QY	1082	CCTGCAGATGTTCTGATGTTTCCAAATCCAGGAACACTAGTTGGAGGCTCCAGCAACAA	1141	Sequence 90, Appl
Db	586	CAAAAAGTGGATGCTGAAGAAAGTCGCTCCTCAAGCTAAAATCGCTGAATTGGAANAATCA	645	Sequence 185, App
QY	1142	GGAAGTAGTATTG	1154	Sequence 327, App

ALIGNMENTS


```
RESULT 8
US-10-767-701-17775
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 17775
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(165223)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1.
US-10-767-701-17775

Query Match      1.8%; Score 39.6; DB 6; Length 447;
Best Local Similarity 60.0%; Pred. No. 1.1;
Matches 66; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy      6 ATCTTTAAAAACAGCTCGCATTAATTATTAGTGAGAGCTTTTATTTTATTTTATATAAT 65
Db      324 ATTTTAAAAAATTATTAATGTTTATTTTATATATATATTTTATTTTATTTTATTTTAT 383

Qy      66 AAAAATAAAAGATTTTATTTTATTTTGGATTTTATGGTTTAAATCCCTATT 115
Db      384 ATTTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 433

RESULT 9
US-10-767-795-178
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 178
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(165223)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1.
US-10-767-795-178

Query Match      1.8%; Score 39.6; DB 6; Length 466;
Best Local Similarity 63.8%; Pred. No. 1.1;
Matches 60; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy      22 TCGCATTAATTAATTAGTGAGAGCTTTTATTTTATTTTATTTTATTAATAAAACTAAAGATTTT 81
Db      265 TTGTTTATTAATTAATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 324

Qy      82 TATTATTTTTCAGTTTATTTTATGGTTAAATCCCTATT 115
Db      325 TATTATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 358

RESULT 10
US-10-796-280-12544
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12544
; LENGTH: 165223
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(165223)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1.
US-10-796-280-12544

Query Match      1.8%; Score 39.2; DB 6; Length 165223;
Best Local Similarity 47.5%; Pred. No. 19;
Matches 116; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy      1022 ATCTTCAAGAGCGGAGCAAAATGGTATACAGCGCTGAGAAAGATCTTAAAAATATCAAA 1081
Db      145793 ATCTCTTTGATGTGAATGAAAAGATGCACCTTGGAAACTGCTAGTAGGCATTTTCAAA 145852

Qy      1082 CCTGCAGATGGTTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAA 1141
Db      145953 TGTAAATGCCATTTTATTAATATATATAAGGATTTATATTTATTTTGGTACTTATAA 145912

Qy      1142 GGAAGTAGTATTTGGTAGTATTCGTTTCATGCTGTAGATGATGCTGAAAAATGAGACC 1201
Db      145913 ATAGTATTATTTTGAATTATTTTCTCATGAAGTTGTTTGTGCAAGTATTGAGAAA 145972

Qy      1202 GCCTCCATTTTGGATGCTGGTTTCGTCAGATGATTCACATGTTCAATACGGAAATCCT 1261
Db      145973 TTATCCCATATTTCTATGTTAAATCTTGACATGGGATTTTACTTCTTTTAGTGACAAAAGT 146032

Qy      1262 GATT 1265
Db      146033 GAAT 146036

RESULT 11
US-10-793-479-31955
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: US/10793479
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/10/793,479
; CURRENT FILING DATE: 2004-03-03
; PRIOR APPLICATION NUMBER: US/09/513,999
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 31955
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-793-479-31955

Query Match      1.8%; Score 38.8; DB 6; Length 335;
Best Local Similarity 54.1%; Pred. No. 1.5;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy      931 CAGTATAAGCAACATAGATTTCAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGC 990
Db      124 CAGTTTAAATACCTTTCTGCAAGCAAAAGAAACAAATCAACAAAGTGAGAGACAAACCCAC 183
```



```
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357
; LENGTH: 2685
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2601)
US-10-784-480-357

Query Match      1.8%; Score 38.2; DB 6; Length 2685;
Best Local Similarity 45.7%; Pred No. 5.5; 248; Indels 12; Gaps 2;
Matches 219; Conservative 0; Mismatches 0;

QY 602 GTTAAAGTTGGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATT 661
Db 1327 GATAGTGAAGGGCGCTAATCTAGGAATTAACGACAAGCTGATGAAGTGGATCTGATT 1386

QY 662 CTGACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCCAAATCT 721
Db 1387 TATAGCGATCTAAGAAACCGTAAACCTTTGATAACATGGCGGCTAAAGGTTATCCATTG 1446

QY 722 GTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTC 781
Db 1447 TTACCAATGGATTTCAAAATGGCGGCGATATTGCCACTATTAACGCCACTAAATGTTGAT 1506

QY 782 CCAGGGAACACGCTGCAATTGCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGACA 841
Db 1507 GCGGACAAAA---TAGCTAGCGATAATCCTATTATTTATGCTTCCATAGAGCTGATATTGCC 1563

QY 842 CAGATAGAGAAAGATGGAAATGCGATTAGGGATGCAATATTTTGACGAGACAGAACGCTA-- 899
Db 1564 AAGCAATACGAACACAGAAAAACCAATTAAGGATAAGATTTTGAAGCTAAATTAGCTAAG 1623

QY 900 -----GTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCA 952
Db 1624 GCTTTAGGTGGCAATAAAAAAGATGACGATAAAGAAAAAGTAAAAAATCCACAGCAGAA 1683

QY 953 GCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCC 1012
Db 1684 GCTAAAGCAGAAAAACAATAAGATAGACAAAGATGTCGCGAAGAACTGCCAAGAAATATCAGT 1743

QY 1013 GACTCTCAAATCTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAA 1071
Db 1744 GAAATCGCTCTTAAGAACAAAAAAGAGAGTGGGGATTTGTAGATGAAATCGTAA 1802
```

Search completed: March 23, 2004, 18:53:26
Job time : 168.201 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:19 ; Search time 5652.79 Seconds

(without alignments)
11389.568 Million cell updates/sec

Title: US-10-608-559-1

Perfect score: 2156

Sequence: 1 ataaatcttttaaacagg.....gatcttcaataataaaga 2156

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_estom:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	52.8	2.4	843	28	BZ643413
C 2	52	2.4	1201	9	AL541034
C 3	50	2.3	1201	13	BX446742
C 4	49	2.3	909	13	BX453105

C 5	48.4	2.2	1191	14	CD500670
C 6	48.2	2.2	533	13	BX353273
C 7	48	2.2	872	14	CK097864
C 8	47.8	2.2	510	28	BH714761
C 9	47	2.2	808	29	CG815585
C 10	46.8	2.2	1201	13	BX356412
C 11	46.6	2.2	359	9	AL840303
C 12	46.6	2.2	1201	9	AL531906
C 13	46.6	2.2	1202	28	CC262481
C 14	46.2	2.1	598	14	CD661508
C 15	46	2.1	427	13	BX403499
C 16	46	2.1	1041	9	AL575932
C 17	46	2.1	1201	9	AL558113
C 18	45.8	2.1	642	13	BX707012
C 19	45.8	2.1	1104	13	BX396222
C 20	45.8	2.1	1201	9	AL561244
C 21	45.6	2.1	967	28	BH184682
C 22	45.6	2.1	967	29	CNS07PKF
C 23	45.6	2.1	1201	13	BX381961
C 24	45.4	2.1	528	12	BJ340195
C 25	45.4	2.1	724	12	BJ344100
C 26	45.4	2.1	819	28	BZ391755
C 27	45.4	2.1	1201	9	AL524496
C 28	45.2	2.1	336	14	CB098017
C 29	45.2	2.1	817	14	CD656959
C 30	45.2	2.1	856	13	BUL75987
C 31	45.2	2.1	890	13	BQ424404
C 32	45.2	2.1	1004	13	BX334499
C 33	45.2	2.1	1072	13	BX359378
C 34	45	2.1	1164	28	CC261365
C 35	44.8	2.1	484	9	AA314486
C 36	44.8	2.1	1124	13	BX436282
C 37	44.8	2.1	1151	10	BE964820
C 38	44.8	2.1	1173	9	AL551058
C 39	44.8	2.1	1183	13	BX379850
C 40	44.6	2.1	259	14	CB722003
C 41	44.6	2.1	1537	12	B1195792
C 42	44.4	2.1	856	28	BH132709
C 43	44.4	2.1	885	13	BX425603
C 44	44.4	2.1	976	13	BX388320
C 45	44.4	2.1	1096	13	BX381337

ALIGNMENTS

RESULT 1
BZ643413/c

LOCUS

DEFINITION OGAOW57TM ZM.0.7.1.5 KB Zea mays genomic clone ZMMBMA011J17,

genomic survey sequence.

ACCESSION BZ643413

VERSION BZ643413.1 GI:28104915

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 843)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Other_GSSs: OGAOW57TC

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

843 bp DNA linear GSS 29-JAN-2003

CD500670 CDA46-D01

BX353273 BX353273

CK097864 UB64CPC02

BH714761 BOMGV71TR

CG815585 SOYFPH83TV

BX356412 BX356412

AL840303 AL840303

AL531906 AL531906

CC262481 CH261-167

CD661508 ETEST6948

BX403499 BX403499

AL575932 AL575932

AL558113 AL558113

BX707012 BX707012

BX396222 BX396222

AL561244 AL561244

BH184682 025_N.14-

AL621633 T3 end of

BX381961 BX381961

BJ340195 BJ340195

BJ344100 BJ344100

BZ391755 EINBH26TR

AL524496 AL524496

CB098017 ku48b04.Y

CD656959 AGENCOURT

BUL75987 AGENCOURT

BQ424404 AGENCOURT

BX334499 BX334499

BX359378 BX359378

CC261365 CH261-72H

AA314486 EST186534

BX436282 BX436282

BE964820 601658475

AL551058 AL551058

BX379850 BX379850

CB722003 jmm603f04

B1195792 602754186

BH132709 ENTNG59TF

BX425603 BX425603

BX388320 BX388320

BX381337 BX381337

```

Class: sheared ends.
Location/Qualifiers
1. .843
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA011J17"
/clone="ZM 0.7 1.5 KB"
/Note="Vector: pBOSK; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      2.4%; Score 52.8; DB 28; Length 843;
Best Local Similarity 47.8%; Pred. No. 1.7;
Matches 153; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 286 TTCTGCAGTGAATGCTCTCATGAGTCTGCAGATAAGCTGGGTATGCTTCTAGTAACAG 345
Db 742 TACTGCTACTACTGCTACTGCTGCTACTGCTACTGCTACTGCTACTGCTACTGCTACT 683

QY 346 CTCGCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGACGCGCACCTACGCC 405
Db 682 TACTGCTACTACTGCTACTGCTGCTACTACTACTACTACTGCTACTGCTACTGCTGTC 623

QY 406 TCCTCCAGCGCTTGTGATGATTATAGACTCAAGCGCAACAGCTTACGATACATCTT 465
Db 622 TACTGCTGCTACTGCTGCTATGCTGCTACTCCAGCTACTACTGCTGCTGCTGCTACTG 563

QY 466 TACCTCAACATCACATGCTGATACATACAGCTGCTTGTGGAGCTCCAGGATGCTGCAC 525
Db 562 TGCTGCTACTGCTACTGCTGCTACTGCTGCTGCTGCTACTGCTGCTACTGCTACTGCT 503

QY 526 TAATAAAGGATACAGCGGCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGGAAAC 585
Db 502 TACTACTACTGCTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTACTGCTACTG 443

QY 586 TAAGATCCGATGCGAGTTA 605
Db 442 TACTACTACTGCTGCTGCTA 423

RESULT 2
AL541034/c      1201 bp      mRNA      linear      EST 12-MAY-2003
LOCUS          AL541034 Homo sapiens PLACENTA Homo sapiens cdna clone CS0DE005YA21
DEFINITION     3-PRIME, mRNA sequence.
ACCESSION      AL541034
VERSION        AL541034.2 GI:30544815
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 951.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cluster.cgi?seq=CS0DE005AAL1NP1&cluster=951.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005AAL1NP1.
Location/Qualifiers
1. .1201

FEATURES
source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YA21"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/Note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match      2.4%; Score 52; DB 9; Length 1201;
Best Local Similarity 42.0%; Pred. No. 2.4;
Matches 63; Conservative 35; Mismatches 52; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACACAGGCTCGCATTAATTATTAGTAGAGCTTTTTTTTTTTTTTT 60
Db 1160 AAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTATTTTATTTTTTTTTTT 1101

QY 61 ATATAAACTTAAAGATTTTTTATTTATTTTTCAGTTTTTATGTTATCTATGTTCC 120
Db 1100 WAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTCGVC 1041

QY 121 AGTCTTATAGACGAACAGACGACACC 150
Db 1040 VMMMMMMVMSGGGMCAMACRACARRAC 1011

RESULT 3
BX446742/c      1201 bp      mRNA      linear      EST 22-MAY-2003
LOCUS          BX446742 Homo sapiens NEUROBLASTOMA Homo sapiens cdna clone
DEFINITION     CL0BB011ZG07 5-PRIME, mRNA sequence.
ACCESSION      BX446742
VERSION        BX446742.1 GI:31025761
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8247.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cluster.cgi?seq=CL0BB011ZG07RPI&cluster=8247.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BB011ZG07RPI.
Location/Qualifiers
1. .1201

FEATURES
source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB011ZG07"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/Note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

```

Query Match 2.3%; Score 49; DB 13; Length 909;

QY 675 GTAACTGACTCTCTTCAGACTCTTACAGGCTGCTCTTCTCCATCTCTAGCAACATA 734
 Db 858 GTGATATTCTTGACATTAGCCGATAAAGTCTGCAAGACAGAGACTGAAGCAGTGCACG 799
 QY 735 ACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCACAGGAAACGC 794
 Db 798 CCACAGATATCAAGCTTTTAAAGACGAGATGAGAGAGATCTAAAGAGGACGATATTC 739
 QY 795 CTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAGATAGAGAAG 854
 Db 738 CCACTATTGTACATTCTACAGATCAAAACCCATCTGTATAAAGTGTATTGAAGAAAGAAG 679
 QY 855 ATGGAAATGCGATTAGGATGATATTTTGCAGGACAGACGCTAGTCGAGCTGTAGAAA 914
 Db 678 AGGAAGAAGAGATGGCGGGAAGGAGAGAGAGAGAAATCAATAGAGAAGAAAGTA 619
 QY 915 ATGCTAAATCTAATAACAGTATAAGCAACATAGA 948
 Db 618 TTGGTAAGAGTGATACGTCTGTACCATATAAAGA 585

RESULT 6
 BX353273/c
 LOCUS
 DEFINITION BX353273 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DC003YH11 5-PRIME, mRNA sequence.

ACCESSION BX353273
 VERSION BX353273.1 GI:30377779
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 533)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DC003CD06Q21.

FEATURES
 source
 1..533
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC003YH11"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dr)
 primer. Five prime end enriched, double-strand cDNA was
 digested with NotI and cloned into the NotI and EcoRV
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 2.2%; Score 48.2; DB 13; Length 533;
 Best Local Similarity 41.4%; Pred. No. 16;
 Matches 48; Conservative 35; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTGAGAGCTTTTATTTTATTTT 60
 Db 431 AAAAAT 372
 QY 61 ATAAATCTTTAAACAGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTT 116
 Db 371 WWW 316

RESULT 7

CK097864
 LOCUS
 DEFINITION UB64CPC02.3P Populus active cambium cDNA library Populus tremula
 cDNA clone UB64CPC02 3', mRNA sequence.
 ACCESSION CK097864
 VERSION CK097864.1 GI:38582189
 KEYWORDS EST.
 SOURCE Populus tremula
 ORGANISM Populus tremula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Malpighiales; Salicaceae; Populus.
 REFERENCE 1 (bases 1 to 872)
 AUTHORS Sterky,F., Bhalerao,R.R., Unneberg,P., Segerman,B., Nilsson,P.,
 Brunner,A.M., Campaa,L., Jonsson-Lindvall,J., Tandre,K.,
 Strauss,S.H., Sundberg,B., Gustafsson,P., Uhlen,M., Bhalerao,R.P.,
 Nilsson,O., Sandberg,G., Karlsson,J., Lundberg,J. and Jansson,S.
 A Populus EST resource for functional genomics
 Unpublished (2003)
 TITLE Unpublished (2003)
 JOURNAL Other_ESTs: UB64CPC02
 COMMENT Contact: Bo Segerman
 Umea Plant Science Center, Department of Plant Physiology
 Umea University
 901 87 Umea, Sweden
 Tel: +46 90 786 5279
 Fax: +46 90 786 6676
 Email: bo.segerman@plantphys.umu.se.

FEATURES
 source
 1..872
 Location/Qualifiers
 /organism="Populus tremula"
 /mol_type="mRNA"
 /db_xref="taxon:113636"
 /clone="UB64CPC02"
 /tissue_type="Active cambium"
 /clone_lib="Populus active cambium cDNA library"

ORIGIN
 Query Match 2.2%; Score 48; DB 14; Length 872;
 Best Local Similarity 50.4%; Pred. No. 17; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 115;
 QY 1261 TGATTCTCAAGCTGCCCAACAGGAGCTCGCAGCACACAGCTAGACGCAAGGCGCTGG 1320
 Db 513 TGCTGCTGCTGCTGCTCAATTGTTGGAGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGC 572
 QY 1321 AGATGACAGTGTCTGCTCAGCGGCTGGCAGATGCTCAGAAAGCTTTAGAAAGGCGCTTAG 1380
 Db 573 TGCTGCAATTGAGGTAGCTGCTGCTGCTGCTGCTGTTTGAAGCTGTGAGAGCTGCTGCTGC 632
 QY 1381 TAAAGCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
 Db 633 TGTGGCTGCTGCTGTTGAACTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 692
 QY 1441 TGTGAGCGCAGGAGTTCCTCCCGCTGCAGCAAGTCTTATAGGTCATCTCTA 1492
 Db 693 TGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744

RESULT 8
 BH714761
 LOCUS
 DEFINITION BOMGV71TR BO 2.3 KB Brassica oleracea genomic clone BOMGV71,
 genomic survey sequence.

ACCESSION BH714761
 VERSION BH714761.1 GI:18808752
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 510)
 AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.


```

/clone_lib="HOMO sapiens PLACENTA CUI 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the NotI and EcoR V

```


Db 103 GCTGCTGCGC 93
|||||

RESULT 15

BX403499
LOCUS BX403499 Homo sapiens PLACENTA mRNA linear EST 13-MAY-2003
DEFINITION 3-PRIME mRNA sequence.
ACCESSION BX403499
VERSION BX403499.1 GI:30635017
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 427)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBAA002ZH05FP1.

FEATURES

Location/Qualifiers
1..427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBAA002ZH05"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 2.1%; Score 46; DB 13; Length 427;
Best Local Similarity 31.5%; Pred. No. 46;
Matches 34; Conservative 47; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGCGCTGCATTAATTATAGTCAGAGCTTTTATTTTATTTT 60
Db 64 WWWWWWWWWWWW WWWWWW WWWWWW WWWWWW WWWWWW WWWWWW WWWWWW 123

QY 61 ATATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 108

Db 124 WTWTAAATWAATWATWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 171

Search completed: March 24, 2004, 05:56:05
Job time : 5664.46 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:24:24 ; Search time 56.3737 Seconds

(without alignments)
3262.843 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MVNPIGPGIDERTPPAD.....SLFIQVLNIGSLGYLIQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep_29Jan04:*

- 1: Genesep1980s:*
- 2: Genesep1990s:*
- 3: Genesep2000s:*
- 4: Genesep2001s:*
- 5: Genesep2002s:*
- 6: Genesep2003as:*
- 7: Genesep2003bs:*
- 8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	3 AAY71954	Aay71954 Chlamydia
2	3187	100.0	651	4 AAU38899	Aau38899 C. pneumo
3	3187	100.0	651	5 ABB94277	Abb94277 Chlamydia
4	3187	100.0	651	5 ABB90532	Abb90532 Chlamydia
5	3187	100.0	651	7 ADD42665	Add42665 Chlamydia
6	2849	89.4	583	3 AAY71955	Aay71955 5'-trunca
7	2208	69.3	490	3 AAY71957	Aay71957 C. pneumo
8	2204	69.2	452	3 AAY71956	Aay71956 3'-trunca
9	2182	68.5	478	2 AAY35358	Aay35358 Chlamydia
10	1000	31.4	212	2 AAY35357	Aay35357 Chlamydia
11	899.5	28.2	647	5 ABB94283	Abb94283 Chlamydia
12	899.5	28.2	647	7 ADD43866	Add43866 Chlamydia
13	894	28.1	660	4 AAU38891	Aau38891 C. tracho
14	894	28.1	660	7 ADD42726	Add42726 Chlamydia
15	671.5	21.1	361	3 AAB13695	Aab13695 Chlamydia
16	671.5	21.1	361	4 AAG83263	Aag83263 Protein e
17	671.5	21.1	361	5 ABB94234	Abb94234 Chlamydia
18	604.5	19.0	331	2 AAY37572	Aay37572 Chlamydia
19	289.5	9.1	350	2 AAY37571	Aay37571 Chlamydia
20	251	7.9	1463	5 AAE20110	Aae20110 Lactobaci
21	235.5	7.4	1178	6 ABB26182	Abb26182 Aspergill
22	234	7.3	1831	6 ABU43109	Abu43109 Protein e
23	234	7.3	1870	6 ABJ19019	Abj19019 Pathogen
24	233	7.3	1822	2 AAR27745	Aar27745 Extracell
25	231.5	7.3	2478	4 AAU37374	Aau37374 Staphyloc

26	231.5	7.3	2478	4 AAU34320	Aau34320 Staphyloc
27	231.5	7.3	2478	6 ABJ19002	Abj19002 Pathogen
28	231.5	7.3	2478	6 ABM71899	Abm71899 Staphyloc
29	223.5	7.0	2481	6 ABU15838	Abu15838 Protein e
30	223.5	7.0	2481	7 ABR62804	Abr62804 Methicill
31	221.5	7.0	971	7 ADC00627	Adc00627 Enterohae
32	218	6.8	603	6 ABU44080	Abu44080 Protein e
33	217.5	6.8	6713	6 ABU15887	Abu15887 Protein e
34	217	6.8	5795	4 AAU37017	Aau37017 Staphyloc
35	215.5	6.8	2016	6 AAE36891	Aae36891 Electreur
36	214.5	6.7	2137	5 ABP39618	Abp39618 Staphyloc
37	213	6.7	2434	4 AAU34339	Aau34339 Staphyloc
38	213	6.7	6281	4 AAU37403	Aau37403 Staphyloc
39	213	6.7	10498	6 ABJ19119	Abj19119 Pathogen
40	210	6.6	1122	6 ABU14859	Abu14859 protein e
41	210	6.6	1123	6 ABU42657	Abu42657 Protein e
42	210	6.6	2271	6 ABU16000	Abu16000 Protein e
43	210	6.6	10182	5 ABP38314	Abp38314 Staphyloc
44	208.5	6.5	4776	6 ABUC2253	Abu02253 S. pneumo
45	208.5	6.5	4776	6 ABU45754	Abu45754 Protein e

ALIGNMENTS

RESULT 1
AAY71954
ID AAY71954 standard; protein; 651 AA.
XX
AC AAY71954;
XX
AC AAY71954;
XX
DT 12-SEP-2003 (revised)
DT 26-MAR-2001 (first entry)
XX
DE Chlamydia pneumoniae 76 kDa full-length protein.
XX

76 kDa protein; bactericidal; diagnosis; prevention; treatment;
pneumonia; upper respiratory tract disease; bronchitis; sinusitis;
acute respiratory disease; cough; sore throat; hoarseness; fever;
vaccine; immunisation.
XX

Chlamydia pneumoniae pneumoniae.
XX

WO2000066739-A2.
XX

09-NOV-2000.
XX

03-MAY-2000; 2000WO-CA000511.
XX

03-MAY-1999; 99US-0132270P.
XX

30-JUN-1999; 99US-0141276P.
XX

(AVET) AVENTIS PASTEUR LTD.
XX

Murdin AD, Coenen RP, Wang J, Dunn P;
XX

WPI: 2000-687542/57.
XX

N-PSDB; AAD02063.
XX

Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
for vaccinating against Chlamydia infections.
XX

Claim 16a; Fig 1; 112pp; English.
XX

The present sequence is Chlamydia pneumoniae full-length 76 kDa protein.
C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
tract disease, bronchitis, sinusitis and acute respiratory disease such
as cough, sore throat, hoarseness, fever; and abnormal chest sounds on
auscultation). C. pneumoniae sequence is also used as vaccines for
immunising humans against diseases caused by C. pneumoniae. (Updated on
12-SEP-2003 to standardise OS field)
XX

off. can

SQ	Sequence 651 AA;	
	Query Match	100.0%; Score 3187; DB 3; Length 651;
	Best Local Similarity	100.0%; Pred. No. 4.6e-194;
	Matches 651; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAQAQIAGAAKPKSKTDSVERWSI	60
DB	1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAQAQIAGAAKPKSKTDSVERWSI	60
QY	61 LRSVAVNALSADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDYKTKQAQAYDT	120
DB	61 LRSVAVNALSADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDYKTKQAQAYDT	120
QY	121 IFTSTSLADIQAALVSLQDAVNIKDTAETETAIAAEWETKNADAQVKGAIITELAKY	180
DB	121 IFTSTSLADIQAALVSLQDAVNIKDTAETETAIAAEWETKNADAQVKGAIITELAKY	180
QY	181 ASDNQAILDSLGLTSPDLLQAALLOSVAANNKAELLKEMQDNPVPGKTPAIAQSLVD	240
DB	181 ASDNQAILDSLGLTSPDLLQAALLOSVAANNKAELLKEMQDNPVPGKTPAIAQSLVD	240
QY	241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATAKTQIAEAQ	300
DB	241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATAKTQIAEAQ	300
QY	301 KKFPDSPILQBAEQMVIAEKDLKNIKPADGSDVNPPTVGGSKQGGSSIGSRVSMML	360
DB	301 KKFPDSPILQBAEQMVIAEKDLKNIKPADGSDVNPPTVGGSKQGGSSIGSRVSMML	360
QY	361 DDAENETASILMSGFROMIHMFENTENPDSQAQOELAAQAAKAAAGDSDSAAALADAQK	420
DB	361 DDAENETASILMSGFROMIHMFENTENPDSQAQOELAAQAAKAAAGDSDSAAALADAQK	420
QY	421 ALEAALGKAGQGGQIILNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSGDYKT	480
DB	421 ALEAALGKAGQGGQIILNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSGDYKT	480
QY	481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPEKTPQALARV	540
DB	481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPEKTPQALARV	540
QY	541 ISGNSRTLGDVYSQVSALOSVWQIIOSNPOANNEEIRQKLTSAVTKPPQFGYPYVOLSND	600
DB	541 ISGNSRTLGDVYSQVSALOSVWQIIOSNPOANNEEIRQKLTSAVTKPPQFGYPYVOLSND	600
QY	601 STQKFTAKLESFAEGSRTAABIKALSPFETNSLFIQCVLVNIGSLYSGYLQ	651
DB	601 STQKFTAKLESFAEGSRTAABIKALSPFETNSLFIQCVLVNIGSLYSGYLQ	651
RESULT 2		
AAU38899	AAU38899 standard; protein; 651 AA.	
XX	AAU38899;	
XX	AAU38899;	
DT	11-SEP-2003 (revised)	
DT	16-JAN-2002 (first entry)	
XX		
DE	C. pneumoniae CT622 homologue CPn0728.	
XX		
KW	Chlamydia; sexually transmitted disease; PID; antibacterial;	
KW	pelvic inflammatory disease; antigen; trachoma; gynecological;	
KW	acute respiratory tract infection; atherosclerosis; male infertility;	
KW	coronary heart disease.	
XX		
OS	Chlamydia pneumoniae.	
XX		
PN	W0200181379.A2.	
XX		
PD	01-NOV-2001	
XX		

23-APR-2001; 2001WO-US013081.
 21-APR-2000; 2000US-0198853P.
 20-JUL-2000; 2000US-0219752P.
 (CORI-) CORIXA CORP.
 Bhatia A, Probst P, Stromberg EJ;
 MPI; 2001-616771/71.
 N-PSDB; AAS57031.
 New polynucleotide for treating Chlamydia infections encodes a polynucleotide containing an immunogenic portion of a Chlamydia antigen.
 Disclosure; Page 161-162; 208pp; English.
 The invention relates to isolated polynucleotide encoding at least a partial Chlamydia protein which is an antigenic fragment, or the components, fragments, homologues and variants, and antibodies raised against the antigenic proteins (or fragments). The nucleic acids, proteins and antibodies are used to diagnose and treat Chlamydia infections (e.g. a sexually transmitted disease, pelvic inflammatory disease (PID), acute respiratory tract infection, trachoma, atherosclerosis and coronary heart disease) in a patient, and in the treatment of male infertility. The compounds of the invention are also useful for detecting the presence of Chlamydia in a patient, and stimulating and/or expanding T cells specific for a Chlamydia protein. The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-2003 to standardise OS field)

Query Match 100.0%; Score 3187; DB 4; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.6e-194;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAQAQIAGAAKPKSKTDSVERWSI 60
 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAQAQIAGAAKPKSKTDSVERWSI 60
 61 LRSVAVNALSADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDYKTKQAQAYDT 120
 61 LRSVAVNALSADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDYKTKQAQAYDT 120
 121 IFTSTSLADIQAALVSLQDAVNIKDTAETETAIAAEWETKNADAQVKGAIITELAKY 180
 121 IFTSTSLADIQAALVSLQDAVNIKDTAETETAIAAEWETKNADAQVKGAIITELAKY 180
 181 ASDNQAILDSLGLTSPDLLQAALLOSVAANNKAELLKEMQDNPVPGKTPAIAQSLVD 240
 181 ASDNQAILDSLGLTSPDLLQAALLOSVAANNKAELLKEMQDNPVPGKTPAIAQSLVD 240
 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATAKTQIAEAQ 300
 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATAKTQIAEAQ 300
 301 KKFPDSPILQBAEQMVIAEKDLKNIKPADGSDVNPPTVGGSKQGGSSIGSRVSMML 360
 301 KKFPDSPILQBAEQMVIAEKDLKNIKPADGSDVNPPTVGGSKQGGSSIGSRVSMML 360
 361 DDAENETASILMSGFROMIHMFENTENPDSQAQOELAAQAAKAAAGDSDSAAALADAQK 420
 361 DDAENETASILMSGFROMIHMFENTENPDSQAQOELAAQAAKAAAGDSDSAAALADAQK 420
 421 ALEAALGKAGQGGQIILNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSGDYKT 480
 421 ALEAALGKAGQGGQIILNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSGDYKT 480
 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPEKTPQALARV 540
 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPEKTPQALARV 540

QY 541 ISGNSRTGLGVYQVSVLSALQVWQIIQSNPOANNEIRKQLTSVTKPPQFGYPVQLSND 600
 Db 541 ISGNSRTGLGVYQVSVLSALQVWQIIQSNPOANNEIRKQLTSVTKPPQFGYPVQLSND 600

QY 601 STQKFIKLESIFAEGRSTAAEIKALSPETNSLFIQQVLVNIIGSLYGYLQ 651
 Db 601 STQKFIKLESIFAEGRSTAAEIKALSPETNSLFIQQVLVNIIGSLYGYLQ 651

RESULT 3
 ABB94277 ID ABB94277 standard; protein; 651 AA.
 XX
 AC ABB94277;
 XX
 DT 29-AUG-2003 (revised)
 DT 05-JUN-2002 (first entry)
 XX
 DE Chlamydia pneumoniae protein sequence SEQ ID NO:405.
 XX
 KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW antibacterial; immunostimulant; immune response;
 KW Chlamydia-specific T-cell response.
 XX
 OS Chlamydia pneumoniae.
 XX
 PN WO200208267-A2.
 XX
 PD 31-JAN-2002.
 XX
 PR 20-JUL-2001; 2001WO-US023121.
 XX
 PR 20-JUL-2000; 2000US-00620412.
 PR 23-APR-2001; 2001US-00841132.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Fling SP, Skeiky YAW, Probst P, Bhatia A;
 XX
 DR WPI; 2002-179901/23.
 XX

Novel compositions comprising Chlamydia Cap1 protein and its use in the treatment of Chlamydia infection.

Disclosure; Page 370-372; 537pp; English.

The present invention describes compositions comprising a Chlamydia Cap1 protein and methods for the diagnosis and therapy of Chlamydia infection. Chlamydia DNA and protein sequences from the present invention can have antibacterial and immunostimulant activities, and can be used in vaccines. Compounds from the present invention can be used for eliciting an immune response, specifically stimulating a Chlamydia-specific T-cell response or inhibiting the development of a Chlamydia infection in an animal. Methods from the present invention can be used for detecting the presence of Chlamydia in a patient; to stimulate and/or expand T cells specific for a Chlamydia protein; and for treatment of a Chlamydia infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent sequences used in the exemplification of the present invention. (Updated on 29-AUG-2003 to standardise OS field)

Sequence 651 AA;

Query Match 100.0%; Score 3187; DB 5; Length 651;

Best Local Similarity 100.0%; Pred. No. 4.6e-194; Mismatches 0; Indels 0; Gaps 0; Matches 651; Conservative 0;

QY 1 MVNPIGPGIDETERTPADLSAQGLEASANKSAEQAQIAEAKPKESKTSVERWSI 60

Db 1 MVNPIGPGIDETERTPADLSAQGLEASANKSAEQAQIAEAKPKESKTSVERWSI 60

QY 61 LRSVNLMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQAYDT 120

Db 61 LRSVNLMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQAYDT 120

QY 121 IFTSTSLADIQAALVSLQDAVTNIKOTAAATDEETAAIAAEMETKNADAVKVGQITELAKY 180
 Db 121 IFTSTSLADIQAALVSLQDAVTNIKOTAAATDEETAAIAAEMETKNADAVKVGQITELAKY 180
 QY 181 ASDNCAILDSLGKLTSPDLLQAALLQSVANNKAAELLKEMQDNPPVPGKTPATAQSLVD 240
 Db 181 ASDNCAILDSLGKLTSPDLLQAALLQSVANNKAAELLKEMQDNPPVPGKTPATAQSLVD 240
 QY 241 QTDATATQIEKDGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEQAQ 300
 Db 241 QTDATATQIEKDGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEQAQ 300
 QY 301 KKFDPSPILQEAQWVIAEKDLKNIKPADGSDVPNFGTTVGGSKQGGSSIGSRVSMML 360
 Db 301 KKFDPSPILQEAQWVIAEKDLKNIKPADGSDVPNFGTTVGGSKQGGSSIGSRVSMML 360
 QY 361 DDAENETASILMSGFROMIHMFNTENPDSQAACQELAAQAAKAAAGDSSAAALADAOK 420
 Db 361 DDAENETASILMSGFROMIHMFNTENPDSQAACQELAAQAAKAAAGDSSAAALADAOK 420
 QY 421 ALEAALGKAGQQQGIILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 Db 421 ALEAALGKAGQQQGIILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 QY 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTFRSVPRAETARGPEKTPQALARY 540
 Db 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTFRSVPRAETARGPEKTPQALARY 540
 QY 541 ISGNSRTGLGVYQVSVLSALQVWQIIQSNPOANNEIRKQLTSVTKPPQFGYPVQLSND 600
 Db 541 ISGNSRTGLGVYQVSVLSALQVWQIIQSNPOANNEIRKQLTSVTKPPQFGYPVQLSND 600
 QY 601 STQKFIKLESIFAEGRSTAAEIKALSPETNSLFIQQVLVNIIGSLYGYLQ 651
 Db 601 STQKFIKLESIFAEGRSTAAEIKALSPETNSLFIQQVLVNIIGSLYGYLQ 651

RESULT 4

ABB90532 ID ABB90532 standard; protein; 651 AA.

XX AC ABB90532;

XX DT 29-AUG-2003 (revised)

XX DT 29-JUL-2002 (first entry)

XX DE Chlamydia pneumoniae cp7033 protein, SEQ ID NO:13.

XX KW Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
 KW human respiratory disease; cardiovascular disease; atherosclerosis;
 KW coronary artery disease; carotid artery stenosis; myocardial infarction;
 KW cerebrovascular disease; aortic aneurysm; claudication; stroke;
 strain CWL029.

XX OS Chlamydia pneumoniae.

XX PN WO200202606-A2.

XX PD 10-JAN-2002.

XX PF 03-JUL-2001; 2001WO-IB001445.

XX PR 03-JUL-2000; 2000GB-00016363.

XX PR 11-JUL-2000; 2000GB-00017047.

XX PR 21-JUL-2000; 2000GB-00017983.

XX PR 07-AUG-2000; 2000GB-00019368.

XX PR 18-AUG-2000; 2000GB-00020440.

XX PR 14-SEP-2000; 2000GB-00022583.

XX PR 10-NOV-2000; 2000GB-00027549.

XX PR 22-DEC-2000; 2000GB-00031706.

(CHIR-) CHIRON SPA.

XX Ratti G, Grandi G;
XX WPI: 2002-154726/20.
XX N-PSDB; ABU91190.
XX Novel Chlamydia pneumoniae protein useful in the manufacture of a
XX medicament for treatment or prevention of infection due to Chlamydia,
XX preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX Claim 1; Page 47; 364pp; English.
XX Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
XX pneumoniae (strain CML029), and ABU91184-ABU91373 represent DNA encoding
XX them. The proteins are predicted to be immunogenic and may therefore be
XX useful in vaccine production and for diagnostic purposes. Chlamydia
XX pneumoniae is a common cause of respiratory disease in humans, and is
XX also involved in the development of cardiovascular diseases such as
XX atherosclerosis, coronary artery disease, carotid artery stenosis,
XX myocardial infarction, cerebrovascular disease, aortic aneurysm,
XX claudication and stroke. The proteins and nucleic acids of the invention
XX may be used in vaccines and pharmaceutical compositions for the
XX prevention or treatment of chlamydial infections, particularly Chlamydia
XX pneumoniae infections. The proteins may also be used in the detection of
XX Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
XX DNA probe assay or blotting techniques for determining Chlamydia
XX pneumoniae gene expression. The present sequence represents a
XX specifically claimed Chlamydia pneumoniae protein of the invention.
XX (Updated on 29-AUG-2003 to standardise OS field)
XX Sequence 651 AA;
Query Match 100.0%; Score 3187; DB 5; Length 651;
Best Local Similarity 100.0%; Pred No. 4.6e-194;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWNPTGPGIDETERTPPADLSAQGLESAANKSAEQAQIAEAKPKESKTSVERWSI 60
DB 1 MWNPTGPGIDETERTPPADLSAQGLESAANKSAEQAQIAEAKPKESKTSVERWSI 60
QY 61 LRSVNAWMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQTAYDT 120
DB 61 LRSVNAWMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQTAYDT 120
QY 121 IFTSTSLADIQAALVSLQDVAVINIKDTATDETAIAEWETKNADAVKVGAIITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDVAVINIKDTATDETAIAEWETKNADAVKVGAIITELAKY 180
QY 181 ASDNQAILDSIGKLTSLFLLQALLOSVANNKAAELKEMQDNPPVPFKTPAIAQSLVD 240
DB 181 ASDNQAILDSIGKLTSLFLLQALLOSVANNKAAELKEMQDNPPVPFKTPAIAQSLVD 240
QY 241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNSNISIDSAKAIAATAKTQIAEAO 300
DB 241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNSNISIDSAKAIAATAKTQIAEAO 300
QY 301 KKFPPSPILQEAQVMIQAEKDLKNIKPADGSDVENPGTTVGGSKQGGSSIGSRVSMML 360
DB 301 KKFPPSPILQEAQVMIQAEKDLKNIKPADGSDVENPGTTVGGSKQGGSSIGSRVSMML 360
QY 361 DDAENETASILMSGFRQIMHNTENPDPSQAQQLAQAARAAKAGDSDSAALADAQK 420
DB 361 DDAENETASILMSGFRQIMHNTENPDPSQAQQLAQAARAAKAGDSDSAALADAQK 420
QY 421 ALEAALGAGQOQILNALGQIASAAVVSAGVPPAAASSIGSSVKOLYKTSKSGSDYKT 480
DB 421 ALEAALGAGQOQILNALGQIASAAVVSAGVPPAAASSIGSSVKOLYKTSKSGSDYKT 480
QY 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPKTKDQALRV 540
DB 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPKTKDQALRV 540
QY 541 ISGNSRTIGDVYSQVLSAQSVNQIIQSNPQANNEIROKLTSAVTKPPQFGYPYQLSND 600

DB 541 ISGNSRTIGDVYSQVLSAQSVNQIIQSNPQANNEIROKLTSAVTKPPQFGYPYQLSND 600
QY 601 STQKFIKLESILFABGSRTAAEIKALSFTNSLFTQQVLVNIIGSLYSGLYQ 651
DB 601 STQKFIKLESILFABGSRTAAEIKALSFTNSLFTQQVLVNIIGSLYSGLYQ 651
RESULT 5
ADD42665
ID ADD42665 standard; protein; 651 AA.
XX AC ADD42665;
XX DT 15-JAN-2004 (first entry)
XX DE Chlamydia pneumoniae antigen protein SEQ ID NO:78.
XX KW Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;
XX KW antiinfectivity; cardiant; antiarteriosclerotic; ophthalmological;
KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
KW tubal obstruction; infertility; male infertility; ocular infection;
KW blindness; acute respiratory tract infection; atherosclerosis;
XX KW coronary heart disease.
XX OS Chlamydia pneumoniae.
XX FN WO2003041560-A2.
XX PD 22-MAY-2003.
XX PF 05-NOV-2002; 2002WO-US035624.
XX PR 06-NOV-2001; 2001US-00012256.
XX PR 08-DEC-2001; 2001US-00007693.
XX PR 15-JUL-2002; 2002US-00197220.
XX PA (CORI) CORIXA CORP.
XX PI Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;
PI Probst P;
XX WPI: 2003-441771/41.
XX N-PSDB: ADD42625.
XX New DNA and proteins comprising a portion of a Chlamydia antigen, useful
XX for diagnosing or treating Chlamydial infections, particularly as
XX vaccines for treating or preventing Chlamydial infections, e.g. pelvic
XX inflammatory disease.
XX Disclosure; SEQ ID NO 78; 275pp; English.
XX The present invention describes compounds and methods for diagnosing and
XX treating Chlamydial infection. Chlamydia polynucleotide and protein
XX sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,
XX antiarteriosclerotic and ophthalmological activities, and can be used in
XX vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
XX compositions or methods from the present invention can be used for the
XX serodiagnosis or treatment of Chlamydial infections, particularly in
XX humans. The polynucleotides, proteins or compositions are particularly
XX useful for stimulating an immune response in a patient, or for
XX stimulating and/or expanding T cells specific for a Chlamydia protein.
XX Specifically, the polynucleotides, proteins or compositions are useful as
XX vaccines for treating or preventing Chlamydial infections including
XX pelvic inflammatory disease (which results in tubal obstruction and
XX infertility in women), male infertility, ocular infection (which may
XX cause blindness), acute respiratory tract infections, atherosclerosis, or
XX coronary heart disease. The present sequence is used in the
XX exemplification of the present invention.
XX Sequence 651 AA;
Query Match 100.0%; Score 3187; DB 7; Length 651;

RESULT 7
 AAY71957
 ID AAY71957 standard; protein; 490 AA.
 AC AAY71957;
 XX
 XX
 DT 26-MAR-2001 (first entry)
 XX
 XX C. pneumoniae 76 kDa protein truncation mutant fusion protein.
 XX
 XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
 KW upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; treatment; fusion protein; truncation mutant;
 XX mutin.
 XX
 XX Chlamydia pneumoniae.
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FT 453..490
 FT Region
 FT /note= "this part of the sequence is unrelated to the C.
 FT pneumoniae 76 kDa protein"
 XX
 XX WO200066739-A2.
 PN
 XX
 XX 09-NOV-2000.
 PD
 XX
 XX 03-MAY-2000; 2000WO-CA000511.
 PF
 XX
 XX 03-MAY-1999; 99US-0132270P.
 PR 30-JUN-1999; 99US-0141276P.
 XX
 XX (AVET) AVENTIS PASTEUR LTD.
 PA
 XX
 XX Murdin AD, Oomen RP, Wang J, Dunn P;
 PI WPI; 2000-687542/67.
 XX N-PSDB; AAD02066.
 DR
 XX
 XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
 PT for vaccinating against Chlamydia infections.
 PT
 XX
 XX Claim 33; Fig 3; 112pp; English.
 PS
 XX The present sequence is a fusion protein comprising a truncated Chlamydia
 CC pneumoniae 76 kDa protein and vector-encoded residues. C. pneumoniae 76
 CC kDa protein is used in the diagnosis, prevention and treatment of C.
 CC pneumoniae infections (e.g. pneumonia, upper respiratory tract disease,
 CC bronchitis, sinusitis and acute respiratory disease such as cough, sore
 CC throat, hoarseness, fever; and abnormal chest sounds on auscultation). C.
 CC pneumoniae sequence is also used as vaccines for immunising humans
 CC against diseases caused by C. pneumoniae
 XX
 XX Sequence 490 AA;
 SQ
 Query Match 69.3%; Score 2208; DB 3; Length 490;
 Best Local Similarity 99.8%; Pred. No. 5.1e-132;
 .Matches 453; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MNPIGPGIDETERTPPADLSAQGLEASAAKSAQAQIAEAQKPKESKTSVERWSI 60
 DB 1 MNPIGPGIDETERTPPADLSAQGLEASAAKSAQAQIAEAQKPKESKTSVERWSI 60
 QY 61 LRSVNMALSLADKLGIASSNSSSTSRADVDSTTATPTPPPTFDYKTKQAQTA YDT 120
 DB 61 LRSVNMALSLADKLGIASSNSSSTSRADVDSTTATPTPPPTFDYKTKQAQTA YDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDBETALAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDBETALAAEWETKNADAVKVGAIITELAKY 180
 QY 181 ASDNQAILDSLGLTSLFDLLQALQLSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVD 240

Db 181 ASDNQAILDSLGLTSLFDLLQALQLSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
 Db 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
 QY 301 KKFPPSPILQEAQWVIAEKDLKNIKPADGSDVNPVPGTTVGSKQOGSSIGSIRVSMML 360
 Db 301 KKFPPSPILQEAQWVIAEKDLKNIKPADGSDVNPVPGTTVGSKQOGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMHFNTENPDQAAQQLAQAARAKAAGDSDAAALADAKQ 420
 Db 361 DDAENETASILMSGFRQMHFNTENPDQAAQQLAQAARAKAAGDSDAAALADAKQ 420
 QY 421 ALEAALGKAGQGGGILNALGQIASAAVVSAGVPP 454
 Db 421 ALEAALGKAGQGGGILNALGQIASAAVVSAGVLP 454
 RESULT 8
 AAY71956
 ID AAY71956 standard; protein; 452 AA.
 XX
 AC AAY71956;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE 3'-truncated Chlamydia pneumoniae 76 kDa protein.
 XX
 KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
 KW upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; treatment; truncation mutant; mutin.
 XX
 OS Chlamydia pneumoniae.
 OS Synthetic.
 XX
 XX WO200066739-A2.
 PN
 XX
 XX 09-NOV-2000.
 PD
 XX
 XX 03-MAY-2000; 2000WO-CA000511.
 PF
 XX
 XX 03-MAY-1999; 99US-0132270P.
 PR 30-JUN-1999; 99US-0141276P.
 XX
 XX (AVET) AVENTIS PASTEUR LTD.
 PA
 XX
 XX Murdin AD, Oomen RP, Wang J, Dunn P;
 PI WPI; 2000-687542/67.
 DR N-PSDB; AAD02065.
 XX
 XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
 PT for vaccinating against Chlamydia infections.
 PT
 XX
 XX Claim 16c; Page 104-106; 112pp; English.
 PS
 XX The present sequence is 3'-truncated Chlamydia pneumoniae 76 kDa protein.
 CC pneumoniae 76 kDa protein is used in the diagnosis, prevention and
 CC treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
 CC tract disease, bronchitis, sinusitis and acute respiratory disease such
 CC as cough, sore throat, hoarseness, fever; and abnormal chest sounds on
 CC auscultation). C. pneumoniae sequence is also used as vaccines for
 CC immunising humans against diseases caused by C. pneumoniae
 XX
 XX Sequence 452 AA;
 SQ
 Query Match 69.2%; Score 2204; DB 3; Length 452;
 Best Local Similarity 100.0%; Pred. No. 8.3e-132;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAQAORLAGAEAKPKSKTDSVERWSI 60
 DB 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAQAORLAGAEAKPKSKTDSVERWSI 60
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRGADVSTTATATPTPPPTFDYKTAQATAYDT 120
 DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRGADVSTTATATPTPPPTFDYKTAQATAYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 180
 QY 181 ASDNQAILDSGLKLTDFDLLOAALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 DB 181 ASDNQAILDSGLKLTDFDLLOAALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
 QY 301 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
 DB 301 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQQLAQAARAKAAGDDSAALADAK 420
 DB 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQQLAQAARAKAAGDDSAALADAK 420
 QY 421 ALEAALGKAGQOQGIINLALGQIASAAV 452
 DB 421 ALEAALGKAGQOQGIINLALGQIASAAV 452

RESULT 9

AAAY35358
 ID AAAY35358 standard; protein; 478 AA.
 XX AC AAAY35358;
 XX DT 17-OCT-2003 (revised)
 DT 13-SEP-1999 (first entry)
 XX Chlamydia pneumoniae surface exposed polypeptide.
 DE Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
 KW neutralising epitope.

XX Chlamydia pneumoniae.

OS WO9927105-A2.
 XX PN 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB001890.

XX PR 21-NOV-1997; 97FR-00014673.
 PR 04-NOV-1998; 98US-0107078P.

XX PA (GEST) GENSET.

XX PI Griffiths R;

XX DR WPI; 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae.

XX PS Page 1156; Disclosure; 1912pp; English.

XX CC AAAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAAY34584) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,

CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -OCT-2003 to standardise OS field)

XX Sequence 478 AA;

Query Match 68.5%; Score 2182; DB 2; Length 478;

Best Local Similarity 100.0%; Pred. No. 2.2e-130;

Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAQAORLAGAEAKPKSKTDSVERWSI 60

DB 3 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAQAORLAGAEAKPKSKTDSVERWSI 62

QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRGADVSTTATATPTPPPTFDYKTAQATAYDT 120

DB 63 LRSVAVNALMSLADKLGIASSNSSSTSRGADVSTTATATPTPPPTFDYKTAQATAYDT 122

QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 180

DB 123 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 182

QY 181 ASDNQAILDSGLKLTDFDLLOAALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240

DB 183 ASDNQAILDSGLKLTDFDLLOAALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 242

QY 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300

DB 243 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 302

QY 301 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360

DB 303 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 362

QY 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQQLAQAARAKAAGDDSAALADAK 420

DB 363 DDAENETASILMSGFRQMIHMFNTENPDSQAQQLAQAARAKAAGDDSAALADAK 422

QY 421 ALEAALGKAGQOQGIINLALGQIASAAV 447

DB 423 ALEAALGKAGQOQGIINLALGQIASAAV 449

RESULT 10

AAAY35357

ID AAAY35357 standard; protein; 212 AA.

XX AC AAAY35357;

XX DT 17-OCT-2003 (revised)

DT 13-SEP-1999 (first entry)

XX Chlamydia pneumoniae transmembrane protein sequence.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;

KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;

KW neutralising epitope.

OS Chlamydia pneumoniae.

XX WO9927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB001890.

XX PR 21-NOV-1997; 97FR-00014673.

PR 04-NOV-1998; 98US-0107078P.

```

PA (GEST ) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-357842/30.
XX
XX Genome sequence of Chlamydia pneumoniae.
XX
XX Page 1155; Disclosure; 1912pp; English.
XX
CC AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
CC and is thought to be a contributing factor in heart disease, sarcoidosis,
CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
CC polypeptides encoded by the open reading frames of the C. pneumoniae
CC vaccines (see AAY34584-Y35879) can be used in immunogenic compositions as
CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
CC be used as immunogenic compositions, especially where the vector directs
CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
CC -Oct-2003 to standardise OS field)
XX
SQ Query Match 31.4%; Score 1000; DB 2; Length 212;
Best Local Similarity 99.5%; Pred. No. 9.7e-56;
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 448 VSAGVPPAAASIGSSVKLYTSTSGDYKTIQISAGVDAYKSNIDAYGRARNDTRDV 507
DQ |||||
DB 9 VSAGVPPAAASIGSSVNLQYTSKSTSGDYKTIQISAGVDAYKSNIDAYGRARNDTRDV 68
QY 508 INNVSTPALTRSPVPRARTARGPEKTDQALRVISGNSRTLGDVYSQVSAQSVMQIIQS 567
DQ |||||
DB 69 INNVSTPALTRSPVPRARTARGPEKTDQALRVISGNSRTLGDVYSQVSAQSVMQIIQS 128
QY 568 NPQANNEIRQKLTSAVTKPPQGYPPYVQLSNDSTQKFTAKLESIFABGSRRTAAEIKALS 627
DQ |||||
DB 129 NPQANNEIRQKLTSAVTKPPQGYPPYVQLSNDSTQKFTAKLESIFABGSRRTAAEIKALS 188
QY 628 FETNSLFTQOVLNIGSLYSGLQ 651
DQ |||||
DB 189 FETNSLFTQOVLNIGSLYSGLQ 212

RESULT 11
ABB94283
ID ABB94283 standard; protein; 647 AA.
XX
AC ABB94283;
XX
XX 05-JUN-2002 (first entry)
XX
DE Chlamydia trachomatis protein sequence SEQ ID NO:436.
XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
KW antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response.
XX
OS Chlamydia trachomatis.
XX
XX WO200208267-A2.
FN
PD 31-JAN-2002.
XX
XX 20-JUL-2001; 2001WO-US023121.
XX
PR 20-JUL-2000; 2000US-00620412.
PR 23-APR-2001; 2001US-00841132.
XX
PA (CORI-) CORIXA CORP.
XX
PI Fling SP, Skeiky YAW, Probst P, Bhatia A;

```

```

XX WPI; 2002-179901/23.
XX
XX Novel compositions comprising Chlamydia CapI protein and its use in the
XX treatment of Chlamydia infection.
XX
XX Claim 34; Page 391-393; 537pp; English.
XX
XX The present invention describes compositions comprising a Chlamydia CapI
XX protein and methods for the diagnosis and therapy of Chlamydia infection.
XX Chlamydia DNA and protein sequences from the present invention can have
XX antibacterial and immunostimulant activities, and can be used in
XX vaccines. Compounds from the present invention can be used for eliciting
XX an immune response, specifically stimulating a Chlamydia-specific T-cell
XX response or inhibiting the development of a Chlamydia infection in an
XX animal. Methods from the present invention can be used for detecting the
XX presence of Chlamydia in a patient; to stimulate and/or expand T cells
XX specific for a Chlamydia protein; and for treatment of a Chlamydia
XX infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
XX sequences used in the exemplification of the present invention
XX
SQ Query Match 28.2%; Score 899.5; DB 5; Length 647;
Best Local Similarity 35.1%; Pred. No. 1e-48;
Matches 234; Conservative 118; Mismatches 262; Indels 53; Gaps 16;

QY 2 VNPICGPDIDETPTPADLSAQLGSAANKSABAQRIAGAEAKPKESKTDTSVERMSIL 61
DQ |||||
DB 15 MNPIINGQI-----ASNSETKESTKESEA-----SPSASSSVSSWGF 52
QY 62 RSANVNLMSLADKLGIASSNSSSTSR8-ADVDTTATATPTPPPTPDYKIQGTAYDT 120
DQ |||||
DB 53 SSALKHLSIRD--AILNKNSSPTDSLQLEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110
QY 121 IFTSTLADIQAALVSLQDAVNIKDTAATDEE-TAJAAEWETKNDAVKYGAQITELAK 179
DQ |||||
DB 111 LENATTLAEYETKADLMAALQDMERLAKOKAEVTRIKEALQEQ----EVIDKLNQLVK 166
QY 180 YASNQAILDSIGKLTSPDLLQALLOSANNKAAELLKEMQNPVVGKTPAIAGSLV 239
DQ |||||
DB 167 LEKONQTLKETLTITTSADQIPAINSOLEINKNSADQIIKDEGQNI---SYEAVLTNAG 223
QY 240 DOTDATATQIEKDGNAIRDAYFAGNAGAVENAKSNNSISNIDSAKAAATATQATAEA 299
DQ |||||
DB 224 EVIKASSEAGHKLQALQSIVDAGDQQAALVQANNSPDNIAATKLLDAAETKNEL 283
QY 300 QKK---FPDSPILQEAQWVIQAEKLNKIKPADGSDVPNFGTTVGGSKQGGSSIGSI-- 354
DQ |||||
DB 284 KQEHGTGLTDSPLVKKAEQISQAQKDIQEIKPS-GSDIPIVGPS-GSAASAGSAGVAGALS 341
QY 355 -----RVSMILLDDAENETASILMSGTFQMIMFNTENPDQQAQOELAAQRAAK---AA 406
DQ |||||
DB 342 SNNSGRISLLDDVDNEMAAIAMQGFMSIEQFNWNNPATAKELQAMEAQLTAMSDQLVG 401
QY 407 GDSAAALADAQXALEALGKAGQOQOILNALQIASAAVWSAGVPPAAASSIGSSVKQ 466
DQ |||||
DB 402 ADGELPAEIQAIDALQAL-KQPSDTGLATAMQVAFAAKVGSGAGTAGTIVQMVKQ 460
QY 467 LYKT--SKSTGSDYKTIQISAGVDAYKSNIDAYGRARNDTRDVINNVSTPALTRSPRAR 524
DQ |||||
DB 461 LYKTAFTSTSSSYAAALSDGYSAYKTLNLSYBSRS-GVQSAISQTANPALSRVSRSRG 519
QY 525 TEAREG-PEKTDQALRVISGNSRTLGDVYSQVSAQSVMQIIQSNPOANNEIRPKLTSA 583
DQ |||||
DB 520 IESQSRADASQRAAETIVRDSQTLGDVYSLQVLDLSMTITVSNPQVQNEINQKLTAS 579
QY 584 VTKPPQFGYPVQLSNDSTQKFIKLESFAEGSRRTAAEIKALSFTENSLFIQOVLNIG 643
DQ |||||
DB 580 ISKAPQFGYPAVQNSADSLQFAAQLEREFVDGERSLAESRENAFRKQPAFIQOVLNIA 639
QY 644 SLYSGYL 650
DQ |||||

```


DR	N-PSDB; AAS65995.
XX	
PT	New polynucleotide for treating Chlamydia infections encodes a
PT	polynucleotides containing an immunogenic portion of a Chlamydia antigen.
XX	
PS	Claim 3; Page 205-207; 208pp; English.
XX	
CC	The invention relates to isolated polynucleotide encoding at least a
CC	partial Chlamydia protein which is an antigenic fragment, or the
CC	complements, fragments, homologues and variants, and antibodies raised
CC	against the antigenic proteins (or fragments). The nucleic acids,
CC	proteins and antibodies are used to diagnose and treat Chlamydia
CC	infections (e.g. a sexually transmitted disease, pelvic inflammatory
CC	disease (PID), acute respiratory tract infection, trachoma,
CC	atherosclerosis and coronary heart disease) in a patient, and in the
CC	treatment of male infertility. The compounds of the invention are also
CC	useful for detecting the presence of Chlamydia in a patient, and
CC	stimulating and/or expanding T cells specific for a Chlamydia protein.
CC	The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-
CC	2003 to standardise OS field)
XX	
XX	Sequence 660 AA:
SO	

Query Match	28.1%	Score	894	DB	4	length	660
Best Local Similarity	34.6%	Pred.No.	2.4e-48				
Matches	233	Conservative	126	Mismatches	262	Indels	52
						Gaps	17
<hr/>							
QY	6	GP	GFIDETERTPADLSAQCLEASANKSAEAFQIAGAEAKPKESKTDSTYERNWLSILSAV	65			
DB	11	GP	SVSNQSMMPINGQIASNSETKSYKA-----SEASP--SASSVSSW\$FLSAK	63			
<hr/>							
QY	66	NALMSIADKLGIASSN\$SSSTBS-ADVDSTTATAPTPPTFDYKTOAQAYDTIFTS	124				
DB	64	NALISLRD--AIINKNSPTDSL\$QLEASTSTSTVTRVAAKYDDEAKSNFDTAK\$GLENA	121				
<hr/>							
QY	125	TSIADIQAALVSLQDAVTNIKDTAATD-----EET-----ATAAEWETKNADAVKVGQAQ	173				
DB	122	KTLAEVETKNADLMAALQDMERLANSPPSNHTEEVNNIKALAEAKDITD-----K	173				
<hr/>							
QY	174	ITELAKYASDNQAILDSIGLKT\$FDLLQ\$AALLQ\$VANNKAAELLKEMQNVPVPGKTPA	233				
DB	174	LNLKLVTLQNK\$SLTEVLKTTDSADQIPALNSQLEINKNSADQIIKDLERQNI---SYEA	230				
<hr/>							
QY	234	IAQSLVDQTDATATQTEKQGNARIDAYFAGNAG\$AVENAKSNNSINISDAKAAIATAK	293				
DB	231	VLTNAQEVIKAS\$EAGIKLQCALQ\$IVDAGQ\$QAAVLQ\$AQONNSPDNIAATKELIDAAE	290				
<hr/>							
QY	294	QTQAEAKK---FPD\$PILQEAQOMVITQAEKDLKNIKPADG\$DVPNPGTTVGG\$KQGG\$S	350				
DB	291	TKVNELKQEH\$TGLT\$PLVKKAEQI\$QAKDICEIKPS-GSDIPIVGPS-G\$ASASGSA	348				
<hr/>							
QY	351	IGSI-----RV\$MLLD\$AENETASILMSGFQMIHMFNTENPD\$QAAQOELAAQARAA	403				
DB	349	AGALK\$NN\$SGRISLL\$DDVDN\$MAATALQGF\$SMIEQFN\$VNP\$PATAKELQAEAOITAM	408				
<hr/>							
QY	404	K---AAGD\$SAAALADAQKALEAALGKAGQOQ\$ILNALGQISAAVVSAGVPPAASSI	460				
DB	409	SDQIVGADGELPAEIQAIKDALLAQL-KQ\$PSADGLATAMQOVAPAAKVGG\$SGTAGTV	467				
<hr/>							
QY	461	G\$SVKQLYKT--\$KSTG\$DYKTOISAGYDAYK\$INDAYGRANDATRDVINNV\$TPALTR	518				
DB	468	QMN\$VQLYKTA\$F\$ST\$SS\$YAAALSDGYSAYKTLNSLY\$SESR\$-GVQ\$AISQ\$TANPALSR	526				
<hr/>							
QY	519	SVPRARTEARG-PEKTDQALRVISGNSR\$TLGDVYSQV\$ALQ\$VMQ\$LIQ\$NPQANNEIR	577				
DB	527	SVSR\$GIESQGR\$ADASQRAEFTIVRDSQTLGDVYSRLQVLDSLMT\$VSNPQANQBEIM	586				
<hr/>							
QY	578	QKLT\$AVTKPQPGY\$YVQV\$LND\$STQ\$FIAKLSLFAEG\$RTAAEIKALS\$FETNSL\$FIQ	637				
DB	587	QKLTASISKAPQGY\$PAVQNSAD\$LQKFAQLERE\$FVDGERSLAE\$S\$ENAFRQKPA\$FIQ	646				
<hr/>							
QY	638	VLWNISL\$Y\$GYL	650				

Db 647 VLWNTASLPSGYL 659

RESULT 14
ADD42726
ID ADD42726 standard; protein; 660 AA.
XX AC ADD42726;
XX XX 15-JAN-2004 (first entry)
XX XX Chlamydia trachomatis antigen protein SEQ ID NO:139.
XX XX Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;
XX XX antiinfertility; cardiant; antiarteriosclerotic; ophthalmological;
XX XX vaccine; gene therapy; immune response; pelvic inflammatory disease;
XX XX tubal obstruction; infertility; male infertility; ocular infection;
XX XX blindness; acute respiratory tract infection; atherosclerosis;
XX XX coronary heart disease.
XX XX Chlamydia trachomatis.
XX OS WO2003041560-A2.
XX XX 22-MAY-2003.
XX XX 05-NOV-2002; 2002WO-US035624.
XX XX 06-NOV-2001; 2001US-00012256.
XX XX 05-DEC-2001; 2001US-00007693.
XX XX 15-JUL-2002; 2002US-00157220.
XX XX (CORI-) CORIXA CORP.
XX XX Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;
XX XX Probst P;
XX XX WPI; 2003-441771/41.
XX XX New DNA and proteins comprising a portion of a Chlamydia antigen, useful
XX XX for diagnosing or treating Chlamydial infections, particularly as
XX XX vaccines for treating or preventing Chlamydial infections, e.g. pelvic
XX XX inflammatory disease.
XX XX Example 4; SEQ ID NO 139; 275pp; English.
XX XX The present invention describes compounds and methods for diagnosing and
XX XX treating Chlamydial infection. Chlamydia polynucleotide and protein
XX XX sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,
XX XX antiarteriosclerotic and ophthalmological activities, and can be used in
XX XX vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
XX XX compositions or methods from the present invention can be used for the
XX XX serodiagnosis or treatment of Chlamydial infections, particularly in
XX XX humans. The polynucleotides, proteins or compositions are particularly
XX XX useful for stimulating an immune response in a patient, or for
XX XX stimulating and/or expanding T cells specific for a Chlamydia protein.
XX XX Specifically, the polynucleotides, proteins or compositions are useful as
XX XX vaccines for treating or preventing Chlamydial infections including
XX XX pelvic inflammatory disease (which results in tubal obstruction and
XX XX infertility in women), male infertility, ocular infection (which may
XX XX cause blindness), acute respiratory tract infections, atherosclerosis, or
XX XX coronary heart disease. The present sequence is used in the
XX XX exemplification of the present invention.
XX XX Sequence 660 AA;

```
Query Match      28.1%; Score 894; DB 7; Length 660;
Best Local Similarity 34.6%; Pred.No.2.e-48;
Matches 233; Conservative 126; Mismatches 262; Indexes 52; Gaps 17

QY      6 GPGC-IDETRTPPADLSAQGLSAANKSAEORIAAEAKPKESKTDVSVERMSILRSAY 65
       |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db     11 GPEGVSSNOSNMPIINGQLASNETKETKA-----SEASP--SASSSVSWSFSSAK 63
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:53:56 ; Search time 17.7616 Seconds

(without alignments)

3525.629 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MVNPIGPGFIDETERTPPAD.....SLFIOQLVNIIGSLGYLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: Piri:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	2 D72042	conserved hypothet
2	3187	100.0	651	2 E86581	CHLPN 76 kDa homol
3	2205	69.2	715	2 I40729	hypothetical 76K p
4	899.5	28.2	647	2 G71490	hypothetical prote
5	232.5	7.3	2055	2 T31110	extracellular matr
6	232	7.3	1822	2 S33441	EF protein - Strept
7	223.5	7.0	2481	2 D90011	FmtB protein [impo
8	221.5	7.0	971	2 B90835	probable tail fibre
9	221.5	7.0	973	2 C85693	probable membrane
10	220	6.9	6713	2 B93921	hypothetical prote
11	210	6.6	1122	2 G64887	probable tail fibre
12	210	6.6	2271	2 F90073	hypothetical prote
13	208.5	6.5	4776	2 E95206	cell wall surface
14	199	6.2	1365	2 T30822	Impl protein - Myc
15	196.5	6.2	1829	2 T24583	hypothetical prote
16	196.5	6.2	2232	2 T34434	hypothetical prote
17	192	6.0	1156	2 T34852	probable secreted
18	190.5	6.0	1063	2 D86731	hypothetical prote
19	189	5.9	2155	2 AD2742	conserved hypothet
20	189	5.9	2155	2 C97523	hypothetical prote
21	187	5.9	1147	2 T35781	hypothetical prote
22	186.5	5.9	1306	2 S25370	MSB2 protein - yea
23	183	5.7	2186	2 H89960	hypothetical prote
24	182.5	5.7	641	2 C82206	hypothetical prote
25	182.5	5.7	1238	2 T03465	methyl-accepting c
26	182.5	5.7	2541	2 S11661	probable exonuclea
27	182	5.7	1269	2 F84730	tal in - mouse
28	178.5	5.6	1566	2 A43607	probable myosin he
29	177	5.6	2090	2 S26058	cell surface anti g
					probable transform

ALIGNMENTS

RESULT 1

D72042
conserved hypothetical protein CP0018 [imported] - Chlamydothila pneumoniae (strains CM
N1Alterate names: chlpn 76 kda homolog 1 (ct622); hypothetical protein CPn0728
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: D72042; D81623

R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72042

A:Molecule type: DNA

A:Residues: 1-651 <ARN>

A:Cross-references: GB:AB001654; GB:AB001363; NID:g4377031; PIDN:AAI18867.1; PID:g43770

A:Experimental source: Strain CWL029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10694935

A:Accession: D81623

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-651 <REA>

A:Cross-references: GB:AB002165; GB:AB002161; NID:g7188948; PIDN:AAF37914.1; PID:g71889

A:Experimental source: strain AR39, HL cells

C:Comment: This sequence was originally identified as homologous to part of a sequence

PIR:H71490).

C:Genetics:

A:Gene: CPn0728; CP0018

Query Match 100.0%; Score 3187; DB 2; Length 651;

Best Local Similarity 100.0%; Pred. No. 1.9e-148;

Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNPIGPGFIDETERTPPADLSAQGLESAANKSAQAQRIAGAAKPKESKTSVERWSI 60

DB 1 MVNPIGPGFIDETERTPPADLSAQGLESAANKSAQAQRIAGAAKPKESKTSVERWSI 60

QY 61 LRSVAVNLMSLADKLGIASSNSSSTSRSDVDSTTATPTPTPTFDYKTKQTAYDT 120

DB 61 LRSVAVNLMSLADKLGIASSNSSSTSRSDVDSTTATPTPTPTFDYKTKQTAYDT 120

QY 121 IFTSTSLADIQAALVSLQDVAVTNIKDTAATDETAIAAEWETKNADAVKVAQITELAKY 180

DB 121 IFTSTSLADIQAALVSLQDVAVTNIKDTAATDETAIAAEWETKNADAVKVAQITELAKY 180

QY 181 ASDNCAILLDSGLKLTSDFLQALQLQSVANNKAAELLKEMODNPVVPKTPAQAQLVD 240

DB 181 ASDNCAILLDSGLKLTSDFLQALQLQSVANNKAAELLKEMODNPVVPKTPAQAQLVD 240

241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAG 300
 Db 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAG 300
 Qy 301 KKFPDSPILQAEQWVIAEAKDLKNIKPADGSDVNPFGTTVGGSKQCGSSIGSRVSMML 360
 Db 301 KKFPDSPILQAEQWVIAEAKDLKNIKPADGSDVNPFGTTVGGSKQCGSSIGSRVSMML 360
 Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 Db 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 Qy 421 ALEALGKAGQCGQILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 Db 421 ALEALGKAGQCGQILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 Qy 481 QISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRARTARGPEKTDQALARV 540
 Db 481 QISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRARTARGPEKTDQALARV 540
 Qy 541 ISGNSRTLGDVYSQVSALQSMQIIQSNPOANNEEIROKLTSAVTKPPQFGYPVQLSND 600
 Db 541 ISGNSRTLGDVYSQVSALQSMQIIQSNPOANNEEIROKLTSAVTKPPQFGYPVQLSND 600
 Qy 601 STQKFIKLESFARGSTAAEIKALSPETNSLFIQQVLVNIIGSLYSGLQ 651
 Db 601 STQKFIKLESFARGSTAAEIKALSPETNSLFIQQVLVNIIGSLYSGLQ 651

RESULT 2

CHLPN 76 kDa homolog 1 (CT622) [imported] - Chlamydia pneumoniae (strain J138)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: E86581
 R:Shirai, M.; Hirakawa, H.; Kimoko, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: E86581
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-651 <STO>
 A:Cross-references: GB:BA000008; NID:G979100; PIDM:BA098935.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: CPJ0728

Query Match 100.0%; Score 3187; DB 2; Length 651;
 Best Local Similarity 100.0%; Pred. No. 1.9e-148;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
 Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
 Qy 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDYKTAQATYDT 120
 Db 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDYKTAQATYDT 120
 Qy 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDETAIAAEWETKNADAVKVGQAQITELAKY 180
 Db 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDETAIAAEWETKNADAVKVGQAQITELAKY 180
 Qy 181 ASDNQAILDSLGLKTSFDLLQALLOSANNKAAELLKEMQDNVVPFGKTPAIQSLVD 240
 Db 181 ASDNQAILDSLGLKTSFDLLQALLOSANNKAAELLKEMQDNVVPFGKTPAIQSLVD 240
 Qy 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAG 300
 Db 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAG 300
 Qy 301 KKFPDSPILQAEQWVIAEAKDLKNIKPADGSDVNPFGTTVGGSKQCGSSIGSRVSMML 360

301 KKFPDSPILQAEQWVIAEAKDLKNIKPADGSDVNPFGTTVGGSKQCGSSIGSRVSMML 360
 Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 Db 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 Qy 421 ALEALGKAGQCGQILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 Db 421 ALEALGKAGQCGQILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 Qy 481 QISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRARTARGPEKTDQALARV 540
 Db 481 QISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRARTARGPEKTDQALARV 540
 Qy 541 ISGNSRTLGDVYSQVSALQSMQIIQSNPOANNEEIROKLTSAVTKPPQFGYPVQLSND 600
 Db 541 ISGNSRTLGDVYSQVSALQSMQIIQSNPOANNEEIROKLTSAVTKPPQFGYPVQLSND 600
 Qy 601 STQKFIKLESFARGSTAAEIKALSPETNSLFIQQVLVNIIGSLYSGLQ 651
 Db 601 STQKFIKLESFARGSTAAEIKALSPETNSLFIQQVLVNIIGSLYSGLQ 651

RESULT 3

140729
 hypothetical 76K protein - Chlamydia pneumoniae (strain AR39)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
 C:Accession: I40729
 R:Perez-Melgosa, M.; Kuo, C.
 Infect. Immun. 62, 880-886, 1994
 A:Title: Isolation and characterization of a gene encoding a Chlamydia pneumoniae 76-k
 A:Reference number: I40729; MUID:94156481; PMID:7509320
 A:Accession: I40729
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-715 <RES>
 A:Cross-references: GB:L23921; NID:G435961; PIDN:AAA23117.1; PID:G435962
 A:Experimental source: strain AR-39
 C:Comment: This is the hypothetical translation of a sequence that was reported as two

Query Match 69.2%; Score 2205; DB 2; Length 715;
 Best Local Similarity 99.6%; Pred. No. 1.9e-100;
 Matches 452; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
 Db 257 LVNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTSVERWSI 316
 Qy 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDYKTAQATYDT 120
 Db 317 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDYKTAQATYDT 376
 Qy 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDETAIAAEWETKNADAVKVGQAQITELAKY 180
 Db 377 IFTSTSLADIQAALVSLQDAVTNIKDTAATDETAIAAEWETKNADAVKVGQAQITELAKY 436
 Qy 181 ASDNQAILDSLGLKTSFDLLQALLOSANNKAAELLKEMQDNVVPFGKTPAIQSLVD 240
 Db 437 ASDNQAILDSLGLKTSFDLLQALLOSANNKAAELLKEMQDNVVPFGKTPAIQSLVD 496
 Qy 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAG 300
 Db 497 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAG 556
 Qy 301 KKFPDSPILQAEQWVIAEAKDLKNIKPADGSDVNPFGTTVGGSKQCGSSIGSRVSMML 360
 Db 557 KKFPDSPILQAEQWVIAEAKDLKNIKPADGSDVNPFGTTVGGSKQCGSSIGSRVSMML 616
 Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 Db 617 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 676

RESULT 6

S33441
EF protein - Streptococcus suis
C:Species: Streptococcus suis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: S33441
R:Smith, H.E.; Reek, F.H.; Vecht, U.; Gielkens, A.L.J.; Smits, M.A.
A:Submitted to the EMBL Data Library, May 1993
A:Description: Repeats in an extracellular protein of weak-pathogenic strains are absent
A:Reference number: S33441
A:Accession: S33441
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1822 <SM1>
A:Cross-references: EMBL:X71880; NID:g298031; PID:CAA50714.1; PID:g298032

Query Match 7.3%; Score 232; DB 2; Length 1822;
Best Local Similarity 21.4%; Pred. No. 0.0014;
Matches 166; Conservative 127; Mismatches 289; Indels 192; Gaps 34;

QY	10	IDTERTPPADLBSAQGLEASAANKSPAQRIAGAENAKPKESKTDVS---BRWILRSAVN	66
Db	1072	IDPNPLTPOE-----KESAGNAVEAAKVATA-ATDKASTPDAVQVEDKGV--AAIN	1122
QY	67	ALMSLADKLGIASSNSSSTSRGAD-----VDSTAT-----APTPTPTFDYKT	112
Db	1123	LIYAKADAGVIAAKLADEIKLEDAQAEKADIDASTMTNEKAIJAKKALQDVDPGKA	1182
QY	1113	Q-----AQYATYDTTFTSTSLADIQAALV-----SLQDAVTNIKDTAADEETAIAAE-WETK	163
Db	1183	ELEDAARVATNETHEATTTEKAKAAELAGBSLTDIGKEARDAVELAKOKELAKEALRTE	1242
QY	164	NADAVKGAQITELAKYA-----SDNQA-----ILDSGLKLTSDFL	199
Db	1243	EEEAATKIVEKLAEDTRKAIEDPNLSDEDQXAEIKKLTDAVAKTLATIRNADKRTQ---	1299
QY	200	LOAALLQSVAANNKAAELIKEMODNPVVPCKTFAIAGSLVDQDATA-TQIEKDGNAIRD	258
Db	1300	-EAEKQAALADLEKAKETQK-IADKAAIDRLTILVKDGELEANTQDANKKIADAARAKE	1357
QY	259	AYFAGONASGA-----VENAKSNNSIS-----NIDSAKAIAA-----	290
Db	1358	AIASNPENLTDAEKKTFTTDAVDAEVAKANDAI SAATSPADVQKEEDAGVAIAEDVLDAAK	1417
QY	291	-TAKTOIAE-----AQKPPSPILQBAEQVIOAEKDLKNIKPADSDVDNPOTTVGGSK	345
Db	1418	QDAKNKIADAAAKAEKAIGSNPNLTDAEKKTFTTDAVDAEVAKANDAI SAATSPADV--QK	1475
QY	346	QOGSSITGSRVSML---LDDAENETASILMSGFRMHMENTENPPSOAQOOLAAQARA	402
Db	1476	EEDAGVAATAEDVLDAAKQDANKKIAKESDPAKSAIDANPNLTDAKESAKKAVDADAKA	1535
QY	403	AKAAGDSDSAAALADAKALEAALGKAGQ-----QOGILNALGQIASRA-----	446
Db	1536	ATDAID--ASTSPVEAQSAEDKGVGSIADVDLDAKQDAKNKIAKEVAAKAEKAI DANPNL	1593
QY	447	-----VVSAGVPPAAASS-----IGSSVKOLYKTSKSGSDYKT	480
Db	1594	SDAEKESAKKAVDADAKATTDADIDASTSPVEAQSAEDKGVGSTRQDVLDAAK---QDAKN	1650
QY	481	QISAGVDAYKSI NDAYGR---ARNDATRDVIN--NVSTPALTFRSVPARTAEARGPE-KT	533
Db	1651	KIAKESDAKSAIDANPNLTDAKESAKKAVDADAKAATDAIDAST--SPVEAQSAEDKG	1708
QY	534	QOMLARVISGNSRTLGDVYSQVS-ALQSVNQIIQSNP-----QANNEEIRQKLTSAVT	585
Db	1709	VGAIAKIIDLAAKQ--DAKNKIAKEAESAKSVIDSNPNLTDAAKEAAKSEIDKAVEEAIV	1766
QY	586	-----KPPQEGPYVQIISNDSTOKFTAKLESUPAESG	617
Db	1767	LINGVRTYQEBEKIKLPMALIKPAKAVTPVVD--PNNLTEKETARTKAFUKENN	1819

RESULT 7

FmtB protein [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: D90011
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, Ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Seki, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; PMID:21311952; PMID:11418146
 A:Accession: D90011
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2481 <KUR>
 A:Cross-references: GB:BA000018; PID:g13701961; PIDN:BAB43253.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: fmtB(mrp)

Query Match	7.0%	Score	223.5	DB 2	Length	2481			
Best local Similarity	22.1%	Pred.	No.	0.0056					
Matches	163	Conservative	129	Mismatches	293	Indels	153	Gaps	31
Qy	4	PIGPGPIDE-----TERTPPADLSAQGLEASAANKSA-EACRIAGAFAKPKESKTD	---	54					
Db	763	PLNPDTTNEVBAIERINAAKVS--GVKAIBATTTAODLERVKNEEIFKIENTTDTSTQT	820						
Qy	55	-----VERWSIURSAVNALMSLADKLGASSNSSSSSTSRSD-----VDS'TATATPT	101						
Db	821	KMDAYKEVROAATARKAQNATVSNATDEEVAEANAADAACTEGLHDIOVVKSOEVADT	880						
Qy	102	PP-----PPFPDYKTKQAQATYDT-----IFPSTSADIQAALVSL---	133						
Db	881	KAKYLDKINAIOTAKVKPAAD---TEVENAYNTRKQEIQNSNAT'TEEKEAAVTELDK	937						
Qy	138	-QDAVNTNIKDTAATD-----EETAIAAEVETKNADAVKVGCAQITELAKYASDNQAILDS	190						
Db	938	KQBARNL-DAANTNSDVTTAKONGIAAINQVQAATTKKSDAK-AETIAKASERKTAIEA	995						
Qy	191	LGKIT-----SFDLLQAALLQSVAANNKA-----AEILLKEMQNPVPVPGKT--PAI	234						
Db	996	MNDSTTEEQAQKDVQAVVTANADIDNATNTVDNAKTTNEATTAATPDANVPAA	1055						
Qy	235	AQSLVDGTATATQIE-KQGNATRDAYFAGN-----ASGAVENAKSNNSI-----	279						
Db	1056	KQAIADKVOQETAIDANNST'TEEKEAAQOVOTEXTAADAATDAHSNVEBAAKNAE	1115						
Qy	280	-----SNIDSAAAIAT---AKTQIAEAQKFPDPSPLQEA--EQWVIOAEKD	322						
Db	1116	IAKIEAIQPATTKDNKAQAIATKANERKTAIAGTDITAEETAAANADVNAVTOAN--	1173						
Qy	323	LKNIKPADG--SDVPNPGTTVGGSKQGGSSIGSIRVMSLLDPAENETASIIILMSGPFIHM	381						
Db	1174	-SNIEAANSQNDVDQAKTTGETSIDQVTPVNNKAT-----ARNEITAILNNKLQETQAT	1227						
Qy	382	FNTEPDSQAQOQLAQAARAAKAGDDSA--ALADQAQKALEAALGKAGQOQG	434						
Db	1228	PDATDEEKOADAE--ANTENGKANQAISAATTNAQVDEAKANAEEAIAINAVTPKVVKKQA	1285						
Qy	435	ILNALGQI-ASAAVWSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKTIQISAGVDAYSIN	493						
Db	1286	AKDEIDQLQATQTVINNDQNA'NKEEKAIIQOQ-----ATAVTDAKNNITA-----	1332						
Qy	494	DAYGRANDATRDVINNVSTPALTRSVPERARTARGEKTDQALARYISGNSRSLTGDVYS	553						
Db	1333	-----ATDDNGVDTAKDAGKNS-IGSTQPAFAVKSNAKNEVDQAVTTQQAIDNTTGATTE	1387						
Qy	554	QVSALQSVMOIISQNPQANNEERQKLTSAVTPPQFGYVQ--LSNDSTQKFTAKULESL	612						
Db	1388	EKNAAKDL--VLKAKGKAYQDILNACTNTNDVTQIKDOAVADIQGITADTTIKQVAK---	1441						

Qy	4	PIGPGPIDE-----TERTPPADLSAQGLEASAANKSA-EACRIAGABAKPKESKTDs----	54
Db	763	PLNPDTTNEEVAEALERINAAKVS--GVKAIBATTTAQDLERVKNEEIFKLENIITDSQT	820
Qy	55	-----VERWSILRSVNMALMSLADKLGITASSNSSSTSRSD-----VDSWTATAPT	101
Db	821	KMDAYKEVROAATARKAQNATVSNATDEEVAENAAVDAATQEGLHDIQVVKSOQEVADT	880
Qy	102	PP-----PPTFDDYKTOAQOATYDT-----LFTSTSLADTQAAALVSL-----	137
Db	881	KAKVLDKINAIOTQAKVKPAAD-----TEVENAYNTRKOEIQNSNASTTEEKEAAATELDAK	937
Qy	138	-QDAVNTNIKDTAATD-----EETAIAAEWETKNADAKVKGAQITETELAKVASDNQAILDS	190
Db	938	KQEARTNL-DAANTNSDVTTAKNGIAAINQVQAATTKKSDAK-AETIAQKASERKTAIEA	995
Qy	191	LGKLT-----SFDLQAAALLQSVANNKA-----AELLKMQNDPVPVPGKT--PAI	234
Db	996	MNDSTTEEQAAKDKVDQAVVTVANADIDNATANTDNDNAKTTNEATIAATTPDANVKPAA	1055
Qy	235	AQSLVDQDTATATQIE-KQGNAIARDAYFAGON-----ASGAVENAKSNNSI-----	279
Db	1056	KQAIADKVQOETAIIDANNNGSTTEEKEAKQOVOTEKTADAAIDAHSNVVEBAKNAE	1115
Qy	280	-----SNIDSAKAAIAT-----AKTQIAEAAQKFPDPSPIQBEA--EQMWIOAEKD	322
Db	1116	IAKIEAIQPATTTKDNKAQAIATKANERKTAIAGTQDITAEETIAAANADVNAVTOAN--	1173
Qy	323	LKNIKPADG--SDVPNPQTTVGGSKQOGSSIGSIRVSMLLDDAENETIASILMSGFROMIHM	381
Db	1174	-SNIEAANSQNDVDQAKTTGETSIDQVTPVWNKAT-----ARNEITAILNNKLOEIQAT	1227
Qy	382	FNTENPDSQAQOELAAQABAAATAGDDSA-------ALADAQKALEALAKGAGQOQG	434
Db	1228	PDATDEKQOADAEE--ANTENGANKAQISAATTNAAQVDEAKANAEEAINAVTPKVVKKQA	1285
Qy	435	ILNALGQI-ASAAVWSAGVPPAAASSITGSSVKQLYKTSKSTGSDYKTOISAGYDAYKSIN	493
Db	1286	AKBEIDOLQATQTNVINNDQNA-FNEEKEAIIQOL-----ATAVTDAKNITA-----	1332
Qy	494	DAYGRARNDATRVINNVSTPALTRSVPRARTARGEKPTQDALARVISGNSRSLTGDVYS	553
Db	1333	-----ATDDNGVDTAKDAGKNS-OSTQPAFAVKSNAKNEVDQAVTTQONQAIIDNTTGATTE	1397
Qy	554	QVSALQSVMOIIOSNPOANNEERQKLTSAVTPPQPGFYVQ--LSNDSTQKFTAKULESL	612
Db	1388	EKNAAKDL--VLKAKGKAYQDILNACTNTDNTVTOIKDOAVADIQGITADTTIKDVAK----	1441

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:24:24 ; Search time 39.1412 Seconds
(without alignments)
3262.843 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204

Sequence: 1 MWNFGPDEITERPPAD.....QGILNALQIASAAVWSAGV 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_29Jan04.*

- 1: geneseq1980s.*
- 2: geneseq1990s.*
- 3: geneseq2000s.*
- 4: geneseq2001s.*
- 5: geneseq2002s.*
- 6: geneseq2003as.*
- 7: geneseq2003bs.*
- 8: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2204	100.0	452	3	AAy71956 3'-trunca
2	2204	100.0	490	3	AAy71957 C. pneumo
3	2204	100.0	651	3	AAy71954 Chlamydia
4	2204	100.0	651	4	Aau38899 C. pneumo
5	2204	100.0	651	5	Abb94277 Chlamydia
6	2204	100.0	651	5	Abb90532 Chlamydia
7	2204	100.0	651	7	Adg42665 Chlamydia
8	2182	99.0	478	2	AAy35358 Chlamydia
9	1866	84.7	583	3	AAy71955 5'-trunca
10	461.5	20.9	647	7	Abb94283 Chlamydia
11	461.5	20.9	647	7	Add43866 Chlamydia
12	453	20.6	660	7	Aau38921 C. tracho
13	453	20.6	660	7	Adg42726 Chlamydia
14	289.5	13.1	350	2	AAy37571 Chlamydia
15	231.5	10.5	361	3	Aabi3695 Chlamydia
16	231.5	10.5	361	4	Ag83263 Chlamydia
17	231.5	10.5	361	5	Abb94234 Chlamydia
18	219	9.9	1463	5	Aae20110 Lactobaci
19	197	8.9	971	7	Adc00627 Enterobae
20	190	8.6	2478	4	AAU37374 Staphyloc
21	190	8.6	2478	4	Aau34320 Staphyloc
22	190	8.6	2478	6	Abj19002 Pathogen
23	190	8.6	2478	6	Abm71899 Staphyloc
24	189.5	8.6	1831	6	Abu43109 Protein e
25	189.5	8.6	1870	6	Abj19019 Pathogen

26	188	8.5	603	6	ABU44080 Protein e
27	186.5	8.5	790	6	ABU47218 Protein e
28	184	8.3	1122	6	ABU14859 Protein e
29	183.5	8.3	1822	2	AAR27745 Extracell
30	181	8.2	2481	6	ABU15838 Protein e
31	181	8.2	2481	7	ABR62804 Methicill
32	180	8.2	2137	5	ABP39618 Staphyloc
33	178	8.1	1178	6	ABJ26182 Aspergill
34	176	8.0	1924	6	ABU43922 Protein e
35	173	7.8	2271	6	ABU16000 Protein e
36	172.5	7.8	812	6	ABU47318 Protein e
37	172	7.8	2016	6	AAE36891 Plectreur
38	172	7.8	2344	4	AAU37120 Staphyloc
39	171.5	7.8	1879	5	AAE20111 Lactobaci
40	170.5	7.7	2368	4	AAU34139 Staphyloc
41	170.5	7.7	2368	4	AAU36796 Staphyloc
42	169.5	7.7	2261	6	ABJ18914 Pathogen
43	168.5	7.6	2398	6	ABU42252 Protein e
44	168.5	7.6	6281	4	AAU37403 Staphyloc
45	168.5	7.6	9535	6	ABM73008 Staphyloc

ALIGNMENTS

RESULT 1
AAy71956
ID AAY71956 standard; protein; 452 AA.
XX AC AAY71956;
XX DT 26-MAR-2001 (first entry)
XX DE 3'-truncated Chlamydia pneumoniae 76 kDa protein.
XX KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
XX KW upper respiratory tract disease; bronchitis; sinusitis;
XX KW acute respiratory disease; cough; sore throat; hoarseness; fever;
XX KW vaccine; immunisation; treatment; truncation mutant; mutein.
XX OS Chlamydia pneumoniae.
XX OS Synthetic.
XX EN WO200066739-A2.
XX PD 09-NOV-2000.
XX PF 03-MAY-2000; 2000WO-CA000511.
XX PR 03-MAY-1999; 99US-0132270P.
XX PR 30-JUN-1999; 99US-0141276P.
XX PA (AVET) AVENTIS PASTEUR LTD.
XX PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX DR WPI; 2000-687542/67.
XX DR N-PSDB; AAD02065.
XX PT Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
XX PT for vaccinating against Chlamydia infections.
XX PS Claim 16c; Page 104-106; 112pp; English.
XX CC The present sequence is 3'-truncated Chlamydia pneumoniae 76 kDa protein.
XX CC C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
XX CC treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
XX CC tract disease, bronchitis, sinusitis and acute respiratory disease such
XX CC as cough, sore throat, hoarseness, fever; and abnormal chest sounds on
XX CC auscultation). C. pneumoniae sequence is also used as vaccines for
XX CC immunising humans against diseases caused by C. pneumoniae
XX SQ Sequence 452 AA;

applicants

```

Query Match      100.0%; Score 2204; DB 3; Length 452;
Best Local Similarity 100.0%; Pred. No. 2.9e-146;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPITGPGIDETERTPPADLSAQGLEASANKSAEARIAGAKPKESKTDVSVERWSI 60
DB 1 MNPITGPGIDETERTPPADLSAQGLEASANKSAEARIAGAKPKESKTDVSVERWSI 60
QY 61 LRSVAVNALMSLADKLGIIASSNSSSTSRSDVSTTATPTPTPTDDYKTOAQYAYDT 120
DB 61 LRSVAVNALMSLADKLGIIASSNSSSTSRSDVSTTATPTPTPTDDYKTOAQYAYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDETAIAAEWETKNADAVKVGQAQITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDETAIAAEWETKNADAVKVGQAQITELAKY 180
QY 181 ASDNQAILDSLGLKTSFDLLQALLOQSVANNKAAELLKEMQDNPVPGKTPAIAQSLVD 240
DB 181 ASDNQAILDSLGLKTSFDLLQALLOQSVANNKAAELLKEMQDNPVPGKTPAIAQSLVD 240
QY 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
DB 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
QY 301 KKFPPSPILQEAQWVIQAEKDLKNIKPADGSDVPNGPTTVGGSKQGGSSIGSIRVSMML 360
DB 301 KKFPPSPILQEAQWVIQAEKDLKNIKPADGSDVPNGPTTVGGSKQGGSSIGSIRVSMML 360
QY 361 DDAENETASILMSGFQMIHNTENPDSQAQOELAAQARAAGDSDSAAALADAKQ 420
DB 361 DDAENETASILMSGFQMIHNTENPDSQAQOELAAQARAAGDSDSAAALADAKQ 420
QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 2
AAV71957
ID AAV71957 standard; protein; 490 AA.
XX
AC AAV71957;
XX
DT 26-MAR-2001 (first entry)
XX
DE C. pneumoniae 76 kDa protein truncation mutant fusion protein.
XX
KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
KW upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; treatment; fusion protein; truncation mutant;
KW mutin.
XX
OS Chlamydothila pneumoniae.
OS Synthetic.
XX
FH Key
FT Region
FT 453..490
FT /note= "This part of the sequence is unrelated to the C.
FT pneumoniae 76 kDa protein"
XX
PN WO200066739-A2.
XX
XX 09-NOV-2000.
XX
XX 03-MAY-2000; 2000WO-CA000511.
XX
XX 03-MAY-1999; 99US-0132270P.
XX 30-JUN-1999; 99US-0141276P.
XX
XX (AVET ) AVENTIS PASTEUR LTD.
XX

```

```

PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX WPI; 2000-687542/67.
DR N-PSDB; AAD02066.
XX
PT Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
PT for vaccinating against Chlamydia infections.
XX
PS Claim 33; Fig 3; 112pp; English.
XX
CC The present sequence is a fusion protein comprising a truncated Chlamydia
CC pneumoniae 76 kDa protein and vector-encoded residues. C. pneumoniae 76
CC kDa protein is used in the diagnosis, prevention and treatment of C.
CC pneumoniae infections (e.g. pneumonia, upper respiratory tract disease,
CC bronchitis, sinusitis and acute respiratory disease such as cough, sore
CC throat, hoarseness, fever; and abnormal chest sounds on auscultation). C.
CC pneumoniae sequence is also used as vaccines for immunising humans
CC against diseases caused by C. pneumoniae
XX
SQ Sequence 490 AA;
Query Match      100.0%; Score 2204; DB 3; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.2e-146;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPITGPGIDETERTPPADLSAQGLEASANKSAEARIAGAKPKESKTDVSVERWSI 60
DB 1 MNPITGPGIDETERTPPADLSAQGLEASANKSAEARIAGAKPKESKTDVSVERWSI 60
QY 61 LRSVAVNALMSLADKLGIIASSNSSSTSRSDVSTTATPTPTPTDDYKTOAQYAYDT 120
DB 61 LRSVAVNALMSLADKLGIIASSNSSSTSRSDVSTTATPTPTPTDDYKTOAQYAYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDETAIAAEWETKNADAVKVGQAQITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDETAIAAEWETKNADAVKVGQAQITELAKY 180
QY 181 ASDNQAILDSLGLKTSFDLLQALLOQSVANNKAAELLKEMQDNPVPGKTPAIAQSLVD 240
DB 181 ASDNQAILDSLGLKTSFDLLQALLOQSVANNKAAELLKEMQDNPVPGKTPAIAQSLVD 240
QY 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
DB 241 QTDATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
QY 301 KKFPPSPILQEAQWVIQAEKDLKNIKPADGSDVPNGPTTVGGSKQGGSSIGSIRVSMML 360
DB 301 KKFPPSPILQEAQWVIQAEKDLKNIKPADGSDVPNGPTTVGGSKQGGSSIGSIRVSMML 360
QY 361 DDAENETASILMSGFQMIHNTENPDSQAQOELAAQARAAGDSDSAAALADAKQ 420
DB 361 DDAENETASILMSGFQMIHNTENPDSQAQOELAAQARAAGDSDSAAALADAKQ 420
QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 3
AAV71954
ID AAV71954 standard; protein; 651 AA.
XX
AC AAV71954;
XX
DT 12-SEP-2003 (revised)
DT 26-MAR-2001 (first entry)
XX
XX Chlamydia pneumoniae 76 kDa full-length protein.
XX
XX 76 kDa protein; bactericidal; diagnosis; prevention; treatment;
XX pneumonia; upper respiratory tract disease; bronchitis; sinusitis;
XX acute respiratory disease; cough; sore throat; hoarseness; fever;
XX vaccine; immunisation.

```

This part of the sequence is unrelated to the C. pneumoniae 76 kDa protein"

WO200066739-A2.

09-NOV-2000.

03-MAY-2000; 2000WO-CA000511.

03-MAY-1999; 99US-0132270P.

30-JUN-1999; 99US-0141276P.

(AVET) AVENTIS PASTEUR LTD.

XX OS Chlamydothila pneumoniae.
XX PN WO200066739-A2.
XX AC
XX PD 09-NOV-2000.
XX PF 03-MAY-2000; 2000WO-CA000511.
XX PR 03-MAY-1999; 99US-0132270P.
XX PR 30-JUN-1999; 99US-0141276P.
XX PA (AVET) AVENTIS PASTEUR LTD.
XX PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX DR WPI; 2000-687542/67.
XX DR N-PSDB; RAND02063.
XX XX
XX PT Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
XX PR for vaccinating against Chlamydia infections.
XX PS Claim 16a; Fig 1; 112pp; English.
XX XX
XX CC The present sequence is Chlamydia pneumoniae full-length 76 kDa protein.
XX CC C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
XX CC treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
XX CC tract disease, bronchitis, sinusitis and acute respiratory disease such
XX CC as cough, sore throat, hoarseness, fever; and abnormal chest sounds on
XX CC auscultation). C. pneumoniae sequence is also used as vaccines for
XX CC immunising humans against diseases caused by C. pneumoniae. (Updated on
XX CC 12-SEP-2003 to standardise OS field)
XX XX
XX SQ Sequence 651 AA;
Query Match 100.0%; Score 2204; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.6e-146; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVNPIGPGIDETERTPPADLSAQGLEASAAKSAEQAQRIAGAEAKPKESKTSVERWSI 60
Db 1 MVNPIGPGIDETERTPPADLSAQGLEASAAKSAEQAQRIAGAEAKPKESKTSVERWSI 60
QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTTDDYKTAQATYDT 120
Db 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTTDDYKTAQATYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVTNPKDTAETFAIAAEWETKNADAVKGAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNPKDTAETFAIAAEWETKNADAVKGAQITELAKY 180
QY 181 ASDNQAILDSLGKLTSPDLLQAALQSVAANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
Db 181 ASDNQAILDSLGKLTSPDLLQAALQSVAANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
QY 241 QTDATATQIEKGNARIDAVFAGNAGAVENAKSNNSINISIDAKAIAATAKTQIAEAQ 300
Db 241 QTDATATQIEKGNARIDAVFAGNAGAVENAKSNNSINISIDAKAIAATAKTQIAEAQ 300
QY 301 KKFPPDSPILOEAEQWVIAEKDLNKPADGSDVNPCTTVGSSKQOQSSISIGSRVSMLL 360
Db 301 KKFPPDSPILOEAEQWVIAEKDLNKPADGSDVNPCTTVGSSKQOQSSISIGSRVSMLL 360
QY 361 DDAENETASILMSGFRMIHMFNTENPDQSAQAQELAAQAPAAKAGDSDSAAALADAQK 420
Db 361 DDAENETASILMSGFRMIHMFNTENPDQSAQAQELAAQAPAAKAGDSDSAAALADAQK 420
QY 421 ALEAALGKAGQOQILNALGQIASAAVVSAGV 452
Db 421 ALEAALGKAGQOQILNALGQIASAAVVSAGV 452

RESULT 4

AAU38899
ID AAU38899 standard; protein; 651 AA.
XX AC
XX AC AAU38899;
XX DT 11-SEP-2003 (revised)
XX DT 16-JAN-2002 (first entry)
XX DE
XX DE C. pneumoniae CT622 homologue CPn0728.
XX KW Chlamydia; sexually transmitted disease; PID; antibacterial;
XX KW pelvic inflammatory disease; antigen; trachoma; gynecological;
XX KW acute respiratory tract infection; atherosclerosis; male infertility;
XX KW coronary heart disease.
XX OS Chlamydothila-pneumoniae.
XX PN WO200181379-A2.
XX PD 01-NOV-2001.
XX PF 23-APR-2001; 2001WO-US013081.
XX PR 21-APR-2000; 2000US-0198953P
XX PR 10-JUL-2000; 2000US-0219752P.
XX PA (CORI-) CORIXA CORP.
XX PI Bhatia A, Probst P, Stromberg EJ;
XX DR WPI; 2001-616771/71.
XX DR N-PSDB; AAS57031.
XX XX
XX PT New polynucleotide for treating Chlamydia infections encodes a
XX PR polynucleotides containing an immunogenic portion of a Chlamydia antigen.
XX PS Disclosure; Page 161-162; 208pp; English.
XX XX
XX CC The invention relates to isolated polynucleotide encoding at least a
XX CC partial Chlamydia protein which is an antigenic fragment, or the
XX CC complements, fragments, homologues and variants, and antibodies raised
XX CC against the antigenic proteins (or fragments). The nucleic acids,
XX CC proteins and antibodies are used to diagnose and treat Chlamydia
XX CC infections (e.g. a sexually transmitted disease, pelvic inflammatory
XX CC disease (PID), acute respiratory tract infection, trachoma,
XX CC atherosclerosis and coronary heart disease) in a patient, and in the
XX CC treatment of male infertility. The compounds of the invention are also
XX CC useful for detecting the presence of Chlamydia in a patient, and
XX CC stimulating and/or expanding T cells specific for a Chlamydia protein.
XX CC The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-
XX CC 2003 to standardise OS field)
XX SQ Sequence 651 AA;
Query Match 100.0%; Score 2204; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.6e-146; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVNPIGPGIDETERTPPADLSAQGLEASAAKSAEQAQRIAGAEAKPKESKTSVERWSI 60
Db 1 MVNPIGPGIDETERTPPADLSAQGLEASAAKSAEQAQRIAGAEAKPKESKTSVERWSI 60
QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTTDDYKTAQATYDT 120
Db 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTTDDYKTAQATYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVTNPKDTAETFAIAAEWETKNADAVKGAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNPKDTAETFAIAAEWETKNADAVKGAQITELAKY 180
QY 181 ASDNQAILDSLGKLTSPDLLQAALQSVAANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
Db 181 ASDNQAILDSLGKLTSPDLLQAALQSVAANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240

QY 241 QTDATATQIEKGNIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 Db 241 QTDATATQIEKGNIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPDPSPILOEAEQWVIQAEKDLKNIKPAGSDVPNGTTVGGSKQGGSSIGSIRVSMML 360
 Db 301 KKFPDPSPILOEAEQWVIQAEKDLKNIKPAGSDVPNGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARAARAKAAGDSDSAAALADAQK 420
 Db 361 DDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARAARAKAAGDSDSAAALADAQK 420
 QY 421 ALFAALGKAGCQGGIILNALGQIASAAVVSAGV 452
 Db 421 ALFAALGKAGCQGGIILNALGQIASAAVVSAGV 452

RESULT 5

ABB94277 standard; protein; 651 AA.
 AC ABB94277;

29-AUG-2003 (revised)
 05-JUN-2002 (first entry)

Chlamydia pneumoniae protein sequence SEQ ID NO:405.

Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
 antibacterial; immunostimulant; immune response;
 Chlamydia-specific T-cell response.

Chlamydia pneumoniae.

WC200208267-A2.

31-JAN-2002.

20-JUL-2001; 2001WO-US023121.

20-JUL-2000; 2000US-00620412.

23-APR-2001; 2001US-00841132.

(CORI-) CORIYA CORP.

Fling SP, Skeiky YAW, Probst P, Bhatia A;

WPI; 2002-179901/23.

Novel compositions comprising Chlamydia Cap1 protein and its use in the treatment of Chlamydia infection.

Disclosure; Page 370-372; 537pp; English.

The present invention describes compositions comprising a Chlamydia Cap1
 protein and methods for the diagnosis and therapy of Chlamydia infection.
 Chlamydia DNA and protein sequences from the present invention can have
 antibacterial and immunostimulant activities, and can be used in
 vaccines. Compounds from the present invention can be used for eliciting
 an immune response, specifically stimulating a Chlamydia-specific T-cell
 response or inhibiting the development of a Chlamydia infection in an
 animal. Methods from the present invention can be used for detecting the
 presence of Chlamydia in a patient; to stimulate and/or expand T cells
 specific for a Chlamydia protein; and for treatment of a Chlamydia
 infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
 sequences used in the exemplification of the present invention. (Updated
 on 29-AUG-2003 to standardise OS field)

Sequence 651 AA;

Query Match

Best Local Similarity 100.0%; Score 2204; DB 5; Length 651;

Pred. No. 4.6e-146;

Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQIAGAEAKPKESKTDVSVERWSI 60
 Db 1 MNNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQIAGAEAKPKESKTDVSVERWSI 60
 QY 61 LRSVNNALMSLADKLGCIASSNSSSTSRSDVSDTTATPTPPPTFFDDYKTAQATAYDT 120
 Db 61 LRSVNNALMSLADKLGCIASSNSSSTSRSDVSDTTATPTPPPTFFDDYKTAQATAYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVNINIKDTAATDEBTAIAAEWETKNADAVKVGAIITELAKY 180
 Db 121 IFTSTSLADIQAALVSLQDAVNINIKDTAATDEBTAIAAEWETKNADAVKVGAIITELAKY 180
 QY 181 ASDNQAILDSLGLKITSFDLLQAALQSVANNKKAELLKEMODNPVVPKTPAIAQSLVD 240
 Db 181 ASDNQAILDSLGLKITSFDLLQAALQSVANNKKAELLKEMODNPVVPKTPAIAQSLVD 240
 QY 241 QTDATATQIEKGNIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 Db 241 QTDATATQIEKGNIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPDPSPILOEAEQWVIQAEKDLKNIKPAGSDVPNGTTVGGSKQGGSSIGSIRVSMML 360
 Db 301 KKFPDPSPILOEAEQWVIQAEKDLKNIKPAGSDVPNGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARAARAKAAGDSDSAAALADAQK 420
 Db 361 DDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARAARAKAAGDSDSAAALADAQK 420
 QY 421 ALFAALGKAGCQGGIILNALGQIASAAVVSAGV 452
 Db 421 ALFAALGKAGCQGGIILNALGQIASAAVVSAGV 452

RESULT 6

ABB90532
 ID ABB90532 standard; protein; 651 AA.

AC ABB90532;

29-AUG-2003 (revised)

29-JUL-2002 (first entry)

Chlamydia pneumoniae cp7033 protein, SEQ ID NO:13.

Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
 human respiratory disease; cardiovascular disease; atherosclerosis;
 coronary artery disease; carotid artery stenosis; myocardial infarction;
 cerebrovascular disease; aortic aneurysm; claudication; stroke;
 strain CWL029.

Chlamydia pneumoniae.

WO200202606-A2.

10-JAN-2002.

03-JUL-2001; 2001WO-IB001445.

03-JUL-2000; 2000GB-00016363.

11-JUL-2000; 2000GB-00017047.

21-JUL-2000; 2000GB-00017983.

07-AUG-2000; 2000GB-00019368.

18-AUG-2000; 2000GB-00020440.

14-SEP-2000; 2000GB-00022583.

10-NOV-2000; 2000GB-00027549.

22-DEC-2000; 2000GB-00031706.

(CHIR-) CHIRON SPA.

Ratti G, Grandi G;

XX

DR WPI; 2002-154726/20.
 DR N-PSDB; ABL91190.
 XX Novel Chlamydia pneumoniae protein useful in the manufacture of a
 PT medicament for treatment or prevention of infection due to Chlamydia,
 PT preferably Chlamydia pneumoniae, and for diagnostic purposes.
 XX
 XX Claim 1; Page 47; 364pp; English.
 XX
 CC Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
 CC pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding
 CC them. The proteins are predicted to be immunogenic and may therefore be
 CC useful in vaccine production and for diagnostic purposes. Chlamydia
 CC pneumoniae is a common cause of respiratory disease in humans, and is
 CC also involved in the development of cardiovascular diseases such as
 CC atherosclerosis, coronary artery disease, carotid artery stenosis,
 CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
 CC claudication and stroke. The proteins and nucleic acids of the invention
 CC may be used in vaccines and pharmaceutical compositions for the
 CC prevention or treatment of chlamydial infections, particularly Chlamydia
 CC pneumoniae infections. The proteins may also be used in the detection of
 CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
 CC DNA probe assay or blotting techniques for determining Chlamydia
 CC pneumoniae gene expression. The present sequence represents a
 CC specifically claimed Chlamydia pneumoniae protein of the invention.
 CC (Updated on 29-AUG-2003 to standardise OS field)
 XX
 XX Sequence 651 AA;

Query Match 100.0%; Score 2204; DB 5; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.6e-146; Indels 0; Gaps 0;
 Matches 452; Conservative 0; Mismatches 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEAOIAGAEAKPKESKTSVERWSI 60
 DB 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEAOIAGAEAKPKESKTSVERWSI 60
 QY 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTDDYKTQATAYDT 120
 DB 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTDDYKTQATAYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 QY 181 ASDNQAILDSLGLKLSFDLLQALQLSVANNKAAELKEMQDNFVPGKTPAIAQSLVD 240
 DB 181 ASDNQAILDSLGLKLSFDLLQALQLSVANNKAAELKEMQDNFVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNNSISNTDSKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNNSISNTDSKAAIATAKTQIAEAQ 300
 QY 301 KKFPPSPILQEAQVIOAEKDKNIKPADGSDVNPPTTGVGSKQGGSSIGSIIRVSMLL 360
 DB 301 KKFPPSPILQEAQVIOAEKDKNIKPADGSDVNPPTTGVGSKQGGSSIGSIIRVSMLL 360
 QY 361 DDAENETASILMSGFRQTHMENTNPDSQAAQQLAAQARAAGDDSDAAALADAAQK 420
 DB 361 DDAENETASILMSGFRQTHMENTNPDSQAAQQLAAQARAAGDDSDAAALADAAQK 420
 QY 421 ALEAALGRAGQGGIILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGRAGQGGIILNALGQIASAAVVSAGV 452

RESULT 7

ADD42665

ID ADD42665 standard; protein; 651 AA.

AC ADD42665;

XX
 XX
 DT 15-JAN-2004 (first entry)

XX Chlamydia pneumoniae antigen protein SEQ ID NO:78.
 DE
 XX Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;
 KW antiinfertility; cardiant; antiarteriosclerotic; ophthalmological;
 KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
 KW tubal obstruction; infertility; male infertility; ocular infection;
 KW blindness; acute respiratory tract infection; atherosclerosis;
 KW coronary heart disease.
 XX
 OS Chlamydia pneumoniae.
 XX
 XX MO2003041560-A2.
 XX 22-MAY-2003.
 XX 05-NOV-2002; 2002WO-US035624.
 XX 06-NOV-2001; 2001US-00012246.
 XX 05-DEC-2001; 2001US-00007493.
 XX 15-JUL-2002; 2002US-00197220.
 XX (CORI-) CORIXA CORP
 XX
 XX Bhatia A, Guérian J, Skeiky YAW, Maisonneuve JL, Barth B;
 PI Probst P;
 XX WPI; 2003-441771/41.
 DR N-PSDB; ADD42825.
 XX
 XX New DNA and proteins comprising a portion of a Chlamydia antigen, useful
 PT for diagnosing or treating Chlamydial infections, particularly as
 PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic
 PT inflammatory disease.
 XX
 PS Disclosure; SEQ ID NO 78; 275pp; English.
 XX
 CC The present invention describes compounds and methods for diagnosing and
 CC treating Chlamydial infection. Chlamydia polynucleotide and protein
 CC sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,
 CC antiarteriosclerotic and ophthalmological activities, and can be used in
 CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
 CC compositions or methods from the present invention can be used for the
 CC serodiagnosis or treatment of Chlamydial infections, particularly in
 CC humans. The polynucleotides, proteins or compositions are particularly
 CC useful for stimulating an immune response in a patient, or for
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC Specifically, the polynucleotides, proteins or compositions are useful as
 CC vaccines for treating or preventing Chlamydial infections including
 CC pelvic inflammatory disease (which results in tubal obstruction and
 CC infertility in women), male infertility, ocular infection (which may
 CC cause blindness), acute respiratory tract infections, atherosclerosis, or
 CC coronary heart disease. The present sequence is used in the
 CC exemplification of the present invention.
 XX
 XX Sequence 651 AA;

Query Match 100.0%; Score 2204; DB 7; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.6e-146;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEAOIAGAEAKPKESKTSVERWSI 60
 DB 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEAOIAGAEAKPKESKTSVERWSI 60
 QY 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTDDYKTQATAYDT 120
 DB 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTDDYKTQATAYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180

QY 181 ASDNQALDSGLKLTSDLLQALLQSVANNKAAELLKEMQDNPVPGKTPAIAQSLVD 240
 DB 181 ASDNQALDSGLKLTSDLLQALLQSVANNKAAELLKEMQDNPVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKTQIAEAQ 300
 DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKTQIAEAQ 300
 QY 301 KKPDPSPILQEAQOMVTOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
 DB 301 KKPDPSPILQEAQOMVTOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
 QY 361 DDENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 DB 361 DDENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 8

AAV35358
 ID AAY35358 standard; protein; 478 AA.

AC AAY35358;

DT 17-OCT-2003 (revised)

DT 13-SEP-1999 (first entry)

XX Chlamydia pneumoniae surface exposed polypeptide.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
 KW neutralising epitope.

OS Chlamydia pneumoniae.

XX W09927105-A2.

XX 03-JUN-1999.

XX 20-NOV-1999; 98WO-IB001890.

XX 21-NOV-1997; 97ER-00014573.

XX 04-NOV-1998; 98US-0107078P.

PA (GEST) GENSET.

XX Griffais R;

XX WPI; 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae.

XX Page 1156; Disclosure; 1912pp; English.

CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAY31990) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,
 CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -OCT-2003 to standardise OS field)

XX Sequence 478 AA;

Query Match

Best Local Similarity 99.0%; Score 2182; DB 2; Length 478;

DR N-PSDB; AAD02064.

Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAANKSAEAQRIAGAEAKPKSKTDSVERWSI 60
 DB 3 MWNPIGPGPIDETERTPPADLSAQGLEASAAANKSAEAQRIAGAEAKPKSKTDSVERWSI 62
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRGADVDDSTTATATPTPPPTFFDDYKTKQAQATYDT 120
 DB 63 LRSVAVNALMSLADKLGIASSNSSSTSRGADVDDSTTATATPTPPPTFFDDYKTKQAQATYDT 122
 QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEBETAABEWETKNADAVKVGQITELAKY 180
 DB 123 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEBETAABEWETKNADAVKVGQITELAKY 182
 QY 181 ASDNQAILDSGLKLTSDLLQALLQSVANNKAAELLKEMQDNPVPGKTPAIAQSLVD 240
 DB 183 ASDNQAILDSGLKLTSDLLQALLQSVANNKAAELLKEMQDNPVPGKTPAIAQSLVD 242
 QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKTQIAEAQ 300
 DB 243 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKTQIAEAQ 302
 QY 301 KKPDPSPILQEAQOMVTOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
 DB 303 KKPDPSPILQEAQOMVTOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 362
 QY 361 DDENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 DB 363 DDENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 422
 QY 421 ALEAALGKAGQOQGIILNALGQIASAAV 447
 DB 423 ALEAALGKAGQOQGIILNALGQIASAAV 449

RESULT 9
 AAY71955
 ID AAY71955 standard; protein; 583 AA.
 AC AAY71955;
 XX 26-MAR-2001 (first entry)
 XX 5'-truncated Chlamydia pneumoniae 76 kDa protein.
 XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
 KW upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; treatment; truncation mutant; mutein.
 XX Chlamydia pneumoniae.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 497 /note= "Encoded by ACT"
 XX W0200066739-A2.
 XX 09-NOV-2000.
 XX 03-MAY-2000; 2000WO-CA000511.
 XX 03-MAY-1999; 99US-0132270P.
 XX 30-JUN-1999; 99US-0141276P.
 XX (AVET) AVENTIS PASTEUR LTD.
 XX Mardin AD, Omen RP, Wang J, Dunn P;
 XX WPI; 2000-687542/67.
 XX DR N-PSDB; AAD02064.

PT Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
 PT for vaccinating against Chlamydia infections.

PS Claim 16b; Page 100-102; 112pp; English.

XX The present sequence is 5'-truncated Chlamydia pneumoniae 76 kDa protein.
 CC C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
 CC treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
 CC tract disease, bronchitis, sinusitis and acute respiratory disease such
 CC as cough, sore throat, hoarseness, fever, and abnormal chest sounds on
 CC auscultation). C. pneumoniae sequence is also used as vaccines for
 CC immunising humans against diseases caused by C. pneumoniae

XX Sequence 583 AA;

Query Match 84.7%; Score 1866; DB 3; Length 583;

Best Local Similarity 100.0%; Pred. No. 2.1e-122;

Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 MSLADKLGIASSSSSTSRSDVDSSTATPTPPPTDDYKTAQATYDTFTSTSLA 128

Db 1 MSLADKLGIASSSSSTSRSDVDSSTATPTPPPTDDYKTAQATYDTFTSTSLA 60

QY 123 DIQAALVSLQDAVNTIKDTAETATAAEWETKNADAVKVGQITELAKYASDQAIL 188

Db 61 DIQAALVSLQDAVNTIKDTAETATAAEWETKNADAVKVGQITELAKYASDQAIL 120

QY 189 DSLGKLTDFDLQALLOSANNKKAELLKEMQNDPVPVPGKTPAIAQSLVDQTDATQ 248

Db 121 DSLGKLTDFDLQALLOSANNKKAELLKEMQNDPVPVPGKTPAIAQSLVDQTDATQ 180

QY 249 LEKQGNALRDAYFAGQNASGAVENAKSNNSISNDSAKAATATKTQTAEAQKPPDSP 308

Db 181 LEKQGNALRDAYFAGQNASGAVENAKSNNSISNDSAKAATATKTQTAEAQKPPDSP 240

QY 309 LQEAQEMVIOAEKDLKNIKPADGSDVPNPPTVGGSKQGGSSIGSIRVSMLLDDAENETA 368

Db 241 LQEAQEMVIOAEKDLKNIKPADGSDVPNPPTVGGSKQGGSSIGSIRVSMLLDDAENETA 300

QY 369 SILMSGFRQMTLMFNTENPDSQAQQELAAQARAAGDSDSAAALADAKALEALGK 428

Db 301 SILMSGFRQMTLMFNTENPDSQAQQELAAQARAAGDSDSAAALADAKALEALGK 360

QY 429 AGQQGGIILNALGQIASAAVVSAGV 452

Db 361 AGQQGGIILNALGQIASAAVVSAGV 384

RESULT 10

ABB94283

ID ABB94283 standard; protein; 647 AA.

XX ABB94283;

AC ABB94283;

XX 05-JUN-2002 (first entry)

DT Chlamydia trachomatis protein sequence SEQ ID NO:436.

DE Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

PA (CORI-) CORIXA CORP.

XX Fling SP, Skeiky YAW, Probst P, Bhatia A;

XX MPI; 2002-179901/23.

DR Novel compositions comprising Chlamydia Cap1 protein and its use in the

XX treatment of Chlamydia infection.

XX Claim 34; Page 391-393; 537pp; English.

XX The present invention describes compositions comprising a Chlamydia Cap1

CC protein and methods for the diagnosis and therapy of Chlamydia infection.

CC Chlamydia DNA and protein sequences from the present invention can have

CC antibacterial and immunostimulant activities, and can be used in

CC vaccines. Compounds from the present invention can be used for eliciting

CC an immune response, specifically stimulating a Chlamydia-specific T-cell

CC response or inhibiting the development of a Chlamydia infection in an

CC animal. Methods from the present invention can be used for detecting the

CC presence of Chlamydia in a patient; to stimulate and/or expand T cells

CC specific for a Chlamydia protein; and for treatment of a Chlamydia

CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent

CC sequences used in the exemplification of the present invention

XX Sequence 647 AA;

QY Query Match 20.9%; Score 461.5; DB 5; Length 647;

Db Best Local Similarity 29.7%; Pred. No. 7.9e-24;

Matches 138; Conservative 86; Mismatches 192; Indels 49; Gaps 13;

QY 2 VNPFGPIDETERTPPADLSAQGLEASAANKSAEAGRIAGAEKPKESKTDSTVERWSIL 61

Db 15 MNPFGPIDETERTPPADLSAQGLEASAANKSAEAGRIAGAEKPKESKTDSTVERWSIL 52

QY 62 RSANVALMSLADKLGIASSSSSTSRPS-ADVDSSTATPTPPPTDDYKTAQATYDT 120

Db 53 SSAGKALISLSD--AILNKSSPTDLSLEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110

QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEE-TAJAAEWETKNADAVKVGQITELAK 179

Db 111 LENATTLAEVETKMDLMAALQDMERLAKQAEVTRIKALQEKQ---EVIDKLNQLVK 166

QY 180 YASDQAILSLGKLTDFDLQALLOSANNKKAELLKEMQNDPVPVPGKTPAIAQSLV 239

Db 167 LEKQNTKLTFTTDSADQIPAINSOLEINKNSADQIIKLEQNI---SYEAVLTNAG 223

QY 240 DQTDATATQTEKQGNALRDAYFAGQNASGAVENAKSNNSISNDSAKAATATKTQTAEA 299

Db 224 EVIKASSEAGIKLQALQSIVDAGDQSQAAVLOAQONNSPDNTAATKLLDAETKYNEL 283

QY 300 QKK---PPDSPILQEAQEMVIOAEKDLKNIKPADGSDVPNPPTVGGSKQGGSSIGSI-- 354

Db 284 KOEHTGLTDSPLVKAEQISQAQKQIETKPS-GSDIPTVGFPS-GSAAGSAGAVGALKS 341

QY 355 ----RVSMLLDDAENETASILMSGFRQMTLMFNTENPDSQAQQELAAQARAAGDSDSAAAL 406

Db 342 SNNSGRISLLEDDVDNEMAAIAMQGFESMIEQFNVPNPATAKELQAMEAQLTAMSDQLVG 401

QY 407 GDSAAALADAKALEALGKAGQGGIILNALGQIASAAVVSAGV 451

Db 402 ADGELPAEIQAIKDALAQAL-KQPSDTGLATAMQVAFRAAKVGG 445

XX RESULT 11

XX ADD43866

XX ADD43866 standard; protein; 647 AA.

XX ADD43866;

XX ADD43866;

XX 15-JAN-2004 (first entry)

XX Chlamydia trachomatis immunogenic protein, SEQ ID NO 161.

11 GPESVSSNOSSNPNPIINGQIASNSTKSTKTA-----SEASP--SASSSVSSVSSFLSSAK 63

66 NALMSIADKLGIASSNSSSSSSRSS-ADVDSTTATAPTPTPTPTDYKYTAQTAIDTIFTS 124

64 NALISLRD--ALNKNSSPTDLSQLEASTSTSVTVAAKDYDEAKSNPDPAKSGLENA 121

125 TSLADIQAALVSLQDAVNIKDTAATD-----EET-----AJAAEWETKNADAVKVGQA 173

122 KTLAEYETKMDLMAALQDMERLANSDPSNNHTEEVNNIKKALEAQKDTID-----K 173

174 ITELAKYASDNQAALDLSGLKLTSPDLLQALLQSVANNKKAELLKEMQDNFVVPKTPA 233

174 LNKLVTLQONKSLTEVTKTDSADQIPAINSOLEINKNSADQIIKDLERQNT---SYEA 230

234 IAQSLVDQDTATATOIEKDGNRAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAATATAK 293

231 VLTNAGEVIKASSEAGIKLGQALQISVDAGDQSGAAVLQAOQNNSPDNTAATKELIDAE 290

294 TQIAEAQKK---FPDPSILQEAQMVIOAEKDLKNIKPADGSDVVPNPTGTVGSGKQGGSS 350

291 TKVNELKQEHGTGLTDSPLVKKAAEQISQAQXDIQEIKFPS-GSDIPIVGPS-GSAASAGSA 348

351 IGSII-----RVSMLLDDAENETASILMSGFROMIHFNTENPDSDQAQOELAAQARA 403

349 AGALKSSNNSGRISLLDDVDNEMRAIALQGFRRSMIEQFVNNNPATAKELQAMEAQLTAM 408

404 K---AAGDSDMAAALADQAQKALEAALGKAGQOQOQILNALGQIASAAVVSAG 451

409 SPQLVGADGELPAETCAIKDALAQL-KQPSADGLATAMGQVAFAPAAKVG 458

RESULT 13

ADD42726

ID ADD42726 standard; protein; 660 AA.

AC ADD42726;

XX

DT 15-JAN-2004 (first entry)

XX

DE Chlamydia trachomatis antigen protein SEQ ID NO:139.

XX

KW Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;

KW antifertility; cardiant; antiarteriosclerotic; ophthalmological;

KW vaccine; gene therapy; immune response; pelvic inflammatory disease;

KW tubal obstruction; infertility; male infertility; ocular infection;

KW blindness; acute respiratory tract infection; atherosclerosis;

KW coronary heart disease.

XX

OS Chlamydia trachomatis.

XX

XX WO2003041560-A2.

XX

PD 22-MAY-2003.

XX

PF 05-NOV-2002; 2002WO-US035624.

XX

XX 06-NOV-2001; 2001US-00012256.

PR 05-DEC-2001; 2001US-00007693.

PR 15-JUL-2002; 2002US-00197220.

XX

XX (CORI-) CORIXA CORP.

PA

XX Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;

PI Probst P;

XX

DR WPI; 2003-441771/41.

XX

PT New DNA and proteins comprising a portion of a Chlamydia antigen, useful

PT for diagnosing or treating Chlamydial infections, particularly as

PT vaccines for treating or preventing Chlamydial infections, e.g., pelvic

PT inflammatory disease.

XX

XX Example 4; SEQ ID NO 139; 275pp; English.

XX The present invention describes compounds and methods for diagnosing and
CC treating Chlamydial infection. Chlamydia polynucleotide and protein
CC sequences have antibiotic, antiinflammatory, antifertility, cardiant,
CC antiarteriosclerotic and ophthalmological activities, and can be used in
CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
CC compositions or methods from the present invention can be used for the
CC serodiagnosis or treatment of Chlamydial infections, particularly in
CC humans. The polynucleotides, proteins or compositions are particularly
CC useful for stimulating an immune response in a patient, or for
CC stimulating and/or expanding T cells specific for a Chlamydia protein.
CC Specifically, the polynucleotides, proteins or compositions are useful as
CC vaccines for treating or preventing Chlamydial infections including
CC pelvic inflammatory disease (which results in tubal obstruction and
CC infertility in women), male infertility, ocular infection (which may
CC cause blindness), acute respiratory tract infections, atherosclerosis, or
CC coronary heart disease. The present sequence is used in the
CC exemplification of the present invention.

Sequence 660 AA:

Query Match	20.6%;	Score 453;	DB 7;	Length 660;
Best Local Similarity	28.9%;	Pred. No. 3.2e-23;		

Qy	6	GGCPIDERTPPADLSAQGLEA	GAANKSAPAOIRIAGAENAKPESKTD	SVVERNSILRS	AV 65	
Db	11	GPESVSSNQSSMNP	IQIASNSETKETKA	-----SEASP--SASSSVSSSFLSSAK	63	
Qy	66	NALMSLADKLG	IASSNSSSSRS-	ADVDSPTATPTPPPTDDYKTAQTAYDIFTS	124	
Db	64	NALISLRD--AL	NKNSSPTUSLQLEASIS	STSTVTWVAAKDDEAKSNEDTAKSGLENA	121	
Qy	125	TSIADIQAALV	SLODAVNIKDTAATD-----EET-----	AIAEWETKNADAVKGAQ	173	
Db	122	KTLAEYETKMAD	LAALQDMERLANS	PSNNHTEVNNIKKALEAQKDTID-----K	173	
Qy	174	ITELAKYASD	NOAILDSIGKLTSPDLLQAALLQSVANNKAEALLKEMQDN	PVPVKTPA	233	
Db	174	LNKLVTLQ	NQNSKLEVLKTTSDAQIPAINSOLEINKNSADQI	IKDLERQNT---SYEA	230	
Qy	234	IAQSLVDQ	TDATATQIEKDGN	AIARDAYFAGONASGAVENAKSNNSISNIDS	AKAALATAK 293	
Db	231	VLTNAGEVI	KASSEAGIKLGALQ	LSIVDAGDOSAAVLQAQQNNSPDNTAA	TWELIDAE 290	
Qy	294	QTQIAQAQK---	FDPSPILQEA	BOMWTQAKSKLKNIKHPADGDSVPNPGT	TVGSKQOOGS 350	
Db	291	TKYNELKQ	EHTGLTDSLPLVKA	EEQISQAQKDQETKPS--GSDIP	IVGPSS-GSAASAGSA 348	
Qy	351	IGSI-----	RVSMLLD	DABENETASTILMSGFR	OMTHMFTENPDSQAQOCELAAQARA 403	
Db	349	AGALAKSS	NNSGRISILLDDVD	NEWAAALQGF	RSMTQOFNVNWPATAKELOAMEQ	OLITAM 408
Qy	404	K---AAGD	DSAAAALADQA	LEALGAKGQOQOQI	LNALQIATASAAVVSAG 451	
Db	409	SDQLVG	ADGELPABTQAIK	ALQAQAL--KOPSADGLATAMGOVAF	AAKYGG 458	

RESULT 14
AAY37571
ID AAY37571 standard; protein; 350 AA.

AC AAY37571;

DT 07-OCT-1999 (first entry)

DE Chlamydia trachomatis surface exposed protein.

KW	Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW	paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW	nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW	bartholinitis; pneumopathy; venereal lymphogranulomatosis

```

OS XX Chlamydia trachomatis.
PN XX WO928475-A2.
XX DT
XX DE
XX PD 10-JUN-1999.
XX PF
XX PR 27-NOV-1998; 98WO-18001939.
XX PR 28-NOV-1997; 97FR-00015041.
XX PR 17-DEC-1997; 97FR-00015034.
XX PR 04-NOV-1998; 98US-0107077P.
XX XX
XX FA (GEST ) GENSET.
XX XX
XX PI Griffais R;
XX XX
XX DR WPI; 1999-371125/31.
XX XX
XX PT Genome sequence of Chlamydia trachomatis.
XX PS
XX PS Disclosure; Page 1226-1227; 1755pp; English.
XX CC
XX CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
XX CC of Chlamydia trachomatis (see AZ01425). The polypeptides can be used as
XX CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
XX CC can also be used to control growth of the microorganism. Chlamydia
XX CC trachomatis is responsible for a large number of diseases, e.g. eye
XX CC diseases such as conventional trachoma, nonendemic trachoma,
XX CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
XX CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
XX CC perinephritis, Bartholinitis; pneumopathy in breast feeding infants; and
XX CC venereal lymphogranulomatosis. The polypeptides of the invention may be
XX CC of use in treating these diseases
XX XX
XX SQ Sequence 350 AA;

Query Match 13.1%; Score 289.5; DB 2; Length 350;
Best Local Similarity 24.9%; Pred. No. 4.1e-12;
Matches 94; Conservative 68; Mismatches 129; Indels 87; Gaps 11;

QY 2 VNPFGPIDEETERTPPADLSAQGLEASANKSAAQRIAGAEKPKESKTDVSVERNSIL 61
DB 23 MNPFGQI-----ANSETKSTKASEA-----SSASSSVSSWSFL 60
QY 62 RSANALMSIADKLGIASNSSSTERS-ADVDSTTATPTPPPTFDYKTOAQATYDT 120
DB 61 SSANALISLRD--AILNKNSSPTSLSQLEASTSTSTVTRVAAKDYDKAKSNFTAKSG 118
QY 121 IFTS-----TSLADIQALVSLQ-----DAVTNIKDTAATDBETAAAEWETK 163
DB 119 LENAKTIAEYETKADLMALQDMEANSDPNSDHTBELNNIKALEAQKDTI----- 170
QY 164 NADAVKVGQITELAKYASDQAILSLGKLTSLQALQLQSVANNKKAELLKEMOD 223
DB 171 -----DKLNKLVILQNKSLTEALTKTDSADQIPAINSLINKSAHIIIEKE 222
QY 224 NFVVPKTPAIAOSLVQDQATATQTEKO-----GNAIRDAYFAGQNASGAVEN 272
DB 223 Q-----ISNYKAVLTIDVEKVIKEFSEAGIKGLQALQISVDAGDQSQAALVQ 268
QY 273 AKSNNSISNTDSAKAIAATAKTOIAEAQ---KKFPDSPILQEAQOMVIOAEKDKLNKIFA 329
DB 269 ARGNSPDNIAATKVLIDAAKTKNELKQHEIADSPVLKKAEBQINQAQDIQTITPS 328
QY 330 DGSDFVNPFGTTVGGSKQ 347
DB 329 -GLDIPVGFSGSXPE 345

RESULT 15
AAB13695
ID AAB13695 standard; protein; 361 AA.
XX
```

```

AC AAB13695;
XX DT
XX DT 02-FEB-2001 (first entry)
XX DE
XX DE Chlamydia sp. protein # 6.
XX XX
XX KW Chlamydial infection; sexually transmitted disease;
XX KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
XX KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;
XX KW coronary heart disease; antibacterial.
XX OS
XX OS Chlamydia sp.
XX PN
XX PN WO200034483-A2.
XX PD
XX PD 15-JUN-2000.
XX PF
XX PF 08-DEC-1999; 99WO-US029012.
XX PR
XX PR 08-DEC-1998; 98US-00208277.
XX PR 08-APR-1999; 99US-00288594.
XX PR 01-OCT-1993; 98US-00410568.
XX PR 22-OCT-1999; 99US-00426571.
XX XX
XX XX (CORI-) CORIXA CORP.
XX XX
XX XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX XX WPI; 2000-431303/37.
XX XX
XX XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX XX comprises immunogenic portion of Chlamydia antigen, which comprises amino
XX XX acid sequence encoded by polynucleotide sequence.
XX PS
XX PS Disclosure; Page 251-253; 256pp; English.
XX CC
XX CC The present invention relates to new nucleic acid sequences and the
XX CC proteins encoded by the nucleic acid sequences. The encoded proteins
XX CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX CC proteins are useful for the serodiagnosis and treatment of Chlamydia
XX CC infection. Chlamidiae are intracellular bacterial pathogens that are
XX CC responsible for a wide variety of human infections. C. trachomatis
XX CC infection is one of the most common sexually transmitted diseases and can
XX CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX CC and infertility. Trachoma due to ocular infection with C. trachomatis is
XX CC the leading cause of preventable blindness worldwide. C. pneumonia is a
XX CC major cause of acute respiratory tract infections in humans and is also
XX CC thought to play a role in the pathogenesis of atherosclerosis and
XX CC coronary heart disease. The present sequence is a protein isolated in the
XX CC present invention
XX XX
XX SQ Sequence 361 AA;

Query Match 10.5%; Score 231.5; DB 3; Length 361;
Best Local Similarity 35.4%; Pred. No. 5e-08;
Matches 57; Conservative 30; Mismatches 61; Indels 13; Gaps 5;

QY 301 KKFPDSPILQEAQOMVIOAEKDKLNKIPADGSDVNPFGTTVGGSKQGGSSIGSI----- 354
DB 2 QBIADSPVLKKAEBQINQAQDIQTITPS-GLDIPVGFPS-GSAASASAGALSSNNS 59
QY 355 -RVSMLLDDAENETASILMSGFROHMFNTENPDQSAQAQELAAQAAPAAK---AAGDSD 410
DB 60 GRISLLDDVDNEMAIAIAQGFQSMIEQFNVNPNATAKELQAMEAQLTAMSDQLVGADGE 119
QY 411 ARAALADAQKALEALGKAGQQQGLNKGALQIASAAVVSAG 451
DB 120 LPAEIQATKDALAQAL-KQPSADGLATAMGQVFAAFAKVG 159

Search completed: March 23, 2004, 19:27:40
Job time : 40.1412 secs
```



```
Db 1155 ASTSLGSLSTISDSTSTSTSDASTSTSESDSTST-SLSESTST---SLSDSTST 1210
Qy 79 SSSSS-----STSRADVDSTTATPTPPPTDDYKTAQATAYDTITFTSLADIQ 131
Db 1211 TSEASSTSESDSTSESTSESTSTSVSD-----STASSTSDASTSVSDSE 1261
Qy 132 AALVSLQAV-TNKKDTAATDEETAABWETKNADAVKVGQAITE-----176
Db 1262 SASSTISESLSTSVSDSTST--STSDASTSTSESDSTSESTSESTSVSDSTGAST 1319
Qy 177 -----LAKYASDNQALLDSGLKLTSLFLLQALLOSANNKAELLKEMQDNVPVPGKT 231
Db 1320 SDSASTSESDASTSLGSLSTSLSDST-----STSDASTSTSESDASTSTSL 1376
Qy 232 PAIAQSLVDQDTATATOTEKDGNAIRDAYFAGNASGAVENAKNN-SISNIDSAKAAIA 290
Db 1377 GSTSTSLSDSTSTSDASTSTSVSDSNASTSLGSLSTSVSDSTSTSDASTST 1436
Qy 291 TAKTQIAEAKKFPDPSILOEAWQVIAEKDLNKEADGSDVNPPTTVGGSKQGGSS 350
Db 1437 ESDSERASTSLGSLSTISDSTST-----TSTSDASTSTSVSESNSTST 1482
Qy 351 IG---SIRVSMLLDAAENETASILMSPROMIHMENTENPDSSQAQOELAAQAAKAG 407
Db 1483 ISESLSTSVSDSTSTSDASTSTST-----VSDSDASTSTSESDASTSTST 1535
Qy 408 DDSAAAA 414
Db 1536 SDSASTS 1542

RESULT 6
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PRN
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match 7.4%; Score 163.5; DB 4; Length 10182;
Best Local Similarity 21.8%; Pred. No. 0.014;
Matches 108; Conservative 66; Mismatches 194; Indels 127; Gaps 20;

Qy 16 TPPADLSAQGLEASAANKSAEQAQRIAGAEA-----KPKES-----KTDV 55
Db 2628 TPQVNSALSKVQAQNKINEAKALLQNKADNSQIVRAKEQLQSQIQAASTDGMTQDST 2687
Qy 56 ERWSILRAVNALMSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTDDY-----110
Db 2688 RNYKNKRAEQAQIQAHSV---INNGDATSQINDAKNTVEQAQR-----DYVEAKS 2737
Qy 111 -----KTAQATAYDTITFTSLADIQAAV-SLODAVTNIKDTATDEETAABWETKN 164
Db 2738 NLRADKSLQAYDTLNRDVLNTDKKPSAVRRYNEAISNR-----KE 2780
Qy 165 ADAVKGAQITELAKYASDNQALLDSGLKLTSLFLLQALLOSANNKAELLKEMQDNVPVPGKT 221
Db 2781 LDTAKADAST-LNTNPSVEQVRDALNKINTVQPKVQAIALLOPKENNELVQAKRL 2839
```

```
Qy 222 QDN-----PVVPGKTP-----AIAQSLVDQDTATATOTEKDGNAIRDAY 260
Db 2840 QDAVNDIPQGTQGTQTTNNYNDKQREARALTSAQRVIDNGDATTQETITSEKSKVEQAM 2899
Qy 261 FAGONASG-----AVENAKNN-----SISNIDSAKAAIATAKTOIAEAOAK 301
Db 2900 QALTNAKSNLRADKNELQATYANKLIENVSTNGKKPASIRQYETAKARI---QNGINDAKN 2956
Qy 302 KFPDPSILOEAWQVIAEKDLNKEADGSDVNPPTTVGGSKQGGSSIRVSMLLD 361
Db 2957 E--AERILGNDNPQSVQTLNKKIKAIQ---PKLTEALNMLQNKENNELVNAKRL 3010
Qy 362 DAENETASILMSPROMIHMENTENPDSSQAQOEL-----AAQARAQAAGDD 409
Db 3011 NAVNDTDP--THGWTQ--ETINNYNAKREAONEIQKANMIINNGDATAQDISSEKSKVE 3066
Qy 410 SAAALADAQKALEA 424
Db 3067 QVLAQONAKNDLRA 3081

RESULT 7
US-08-216-894-8
; Sequence 8, Application US/08216894
; Patent No. 5876734
; GENERAL INFORMATION:
; APPLICANT: Kirchhoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,894
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 643 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-216-894-8

Query Match 7.0%; Score 154.5; DB 2; Length 643;
Best Local Similarity 21.5%; Pred. No. 0.0014;
Matches 99; Conservative 59; Mismatches 168; Indels 135; Gaps 17;

Qy 7 PGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAEPKSKTDSVSRVSMILSAVN 66
Db 216 PPKSDILVPRGSPQL--QQAENNTITSKKEMTKLREKVKKAEXEKLDAINRAKLEBERN 273
Qy 67 ALMSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTDDYKTAQATAYDTITFTST 126
Db 274 QAYKAAHK---AEEKAKTFQRLITFESINLNKKRP-----307
```


QY 127 LADIAQALVSLQDAVTN-----IKDTAATDE-----ETAIAAEWETKNADAVKVGQAI 174
 DB 308 -----NDVSNRDKKKNSETAKTDEVEKQRAAEAAKAVETEKQRA-----AEA 350
 QY 175 TELAKYASDNQAILDSLGLTSFDLLQALLOSANNKAAELLENKQDNPNVPGKTPAI 234
 DB 351 TKVAE-----AEKKAEEAAKAVETEKQRAAEATKV 381
 QY 235 A-----QSLVDQDTATATQIEKDGNAIRDYFAGQNASGAVENAKSN-----NSISNID 283
 DB 382 AEAQKQAAEAARAKAVETEKQRAAEATKVAAEKQRAAEAMKVAEAEKQAAEAATKVAAE 441
 QY 284 SAKAAIATAKTQIAEAOAK-KFPDSPILQEAQOM-----VIOAEKDLKNIKPADGSDVP 335
 DB 442 KQKAAEA---TKVAEAEKQKAAEAATKVAAEAEKQAAEAATKVAAEAEKQ---KAAEAATKVA 494
 QY 336 NPGTTVGGSKQGGSSIGSIRVSMILLDDAENETASILMSGFRQMIHMFNTENPDSQAAQOE 395
 DB 495 E-----AEKQKAAEAATKVA-----EAEKQKAAEAATK-----VAEAEKQKAAEAAT 533
 QY 396 LAAQARAAGDSDGAAALADADAQKALEAALGKAGQOQOIL 436
 DB 534 KVAEAEKQKAA--EATKVAAEAEKQKAAEAARAKAVESQKQRF 572

RESULT 8

US-09-115-746-8
 ; Sequence 8, Application US/09115746
 ; Patent No. 6228601
 ; GENERAL INFORMATION:
 ; APPLICANT: Kirchhoff, Louis V.
 ; APPLICANT: Otsu, Keiko
 ; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
 ; TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/115,746
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/216,894
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 643 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-115-746-8

Query Match 7.0%; Score 154.5; DB 3; Length 643;
 Best Local Similarity 21.5%; Pred. No. 0.0014;
 Matches 99; Conservative 59; Mismatches 168; Indels 135; Gaps 17;

QY 7 PGPIDETERTPADLSAQGLEASAAKSAEQAORITAGAEAKPEKSKTDSVERWSILRSVAV 66
 DB 216 PPKSCLVPRGSPSQL--QQAENITNGKEMTKLREKVKAEEKLDAINRAATKLEBERN 273
 QY 67 ALMSLAUKLIGTASSNSSSTSRSDVSTTATATPPPTDDYKTAQAYDTIFISTS 126
 DB 274 QAYKAAHK--AEBEKARTFORLITFESININLKKRP----- 307
 QY 127 LADIAQALVSLQDAVTN-----IKDTAATDE-----ETAIAAEWETKNADAVKVGQAI 174
 DB 308 -----NDVSNRDKKKNSETAKTDEVEKQRAAEAAKAVETEKQRA-----AEA 350
 QY 175 TELAKYASDNQAILDSLGLTSFDLLQALLOSANNKAAELLENKQDNPNVPGKTPAI 234
 DB 351 TKVAE-----AEKKAEEAAKAVETEKQRAAEATKV 381
 QY 235 A-----QSLVDQDTATATQIEKDGNAIRDYFAGQNASGAVENAKSN-----NSISNID 283
 DB 382 AEAQKQAAEAARAKAVETEKQRAAEATKVAAEKQRAAEAMKVAEAEKQAAEAATKVAAE 441
 QY 284 SAKAAIATAKTQIAEAOAK-KFPDSPILQEAQOM-----VIOAEKDLKNIKPADGSDVP 335
 DB 442 KQKAAEA---TKVAEAEKQKAAEAATKVAAEAEKQAAEAATKVAAEAEKQ---KAAEAATKVA 494
 QY 336 NPGTTVGGSKQGGSSIGSIRVSMILLDDAENETASILMSGFRQMIHMFNTENPDSQAAQOE 395
 DB 495 E-----AEKQKAAEAATKVA-----EAEKQKAAEAATK-----VAEAEKQKAAEAAT 533
 QY 396 LAAQARAAGDSDGAAALADADAQKALEAALGKAGQOQOIL 436
 DB 534 KVAEAEKQKAA--EATKVAAEAEKQKAAEAARAKAVESQKQRF 572

RESULT 9
 US-08-894-017-23
 ; Sequence 23, Application US/08894017
 ; Patent No. 6024958
 ; GENERAL INFORMATION:
 ; APPLICANT: Lehner, Thomas
 ; APPLICANT: Kelly, Charles
 ; TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE
 ; TITLE OF INVENTION: OF COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/II
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006-1888
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: Fast-SEQ for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/894,017
 ; FILING DATE: 31-JUL-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB96/00207
 ; FILING DATE: 31-JAN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murashige, Kate H
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 25150-20067.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-887-1500
 ; TELEFAX: 202-822-0168
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1561 amino acids

QY 311 -EAEQWVIAEQKLNKTPADGSDVPNPGTTVGGSKQGSIGSIRVSMLLDRAENETAS 369
Db 285 ADYEAQLAQYKDLIAAQSGSNATNEADYQAKAAVEQELARV-----QAANAAK 334
QY 370 ILMSGFQMIHMENTENPDSQAQOELAAQARAQAAGDSDSAAALADAAQALEAA 425
Db 335 ---QAYEQALAAANTAKN--AQITANEAIQORNAQAKA--NYEAKLQAYQKDLAAA 383

RESULT 14
US-08-216-894-2
; Sequence 2, Application US/08216894
; Patent No. 5876734
; GENERAL INFORMATION:
; APPLICANT: Kirchhoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,894
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 85326/102/DELO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-216-894-2

Query Match 6.5%; Score 142.5; DB 2; Length 564;
Best Local Similarity 21.4%; Pred. No. 0.0095;
Matches 97; Conservative 55; Mismatches 164; Indels 137; Gaps 17;

QY 7 PGPIDETERTPPADLSAQGLEASAANKSABAOIAGAEAKPKESKTDVSVERWSILRSVAVN 66
Db 216 PPKSDLVPRGSPQL--QQAENNIITNSKKEMTKLREKVKAEKELDAINRAKLEERN 273
QY 67 ALMSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTDDYKTOAQATYDTFTSTS 126
Db 274 QAYKAAHK---ABEEKAKTFQRLITFSEINILKRP----- 307
QY 127 LADIQAAALVSLQDAVTN-----IKDTATDE-----ETAIAAEWETKNADAVKGAQI 174
Db 308 -----NDKVSNRDKKKNSETAKTDEVEKQRAAEAAKAVETEKORA-----AEA 350
QY 175 TELAKYASDNQAILDSGLKLTSLDQLAALLQSVANNKAAELIKEMQDNPNVVPKTPAI 234
Db 351 TKVAE-----AEKKAAAEAAKAVETEKORAAEATKV 381
QY 235 A-----QSLVQDTATATQIEKGNADKAVFAGNAGAVENAKSNISNIDSAKAAIA 280
Db 382 AEAEKQKAAEAAKAVETEKORAAEATKVAAEAEKQRAAEAMK-----VAEAEKQKAAEA 434

QY 291 TAKTOIAEAKQ-KFPDPSPIQAEQM-----VIOAEKDLKNIKPADGSDVPNPGTTVG 342
Db 435 ---TKVAEAEKQKAAEATKVAAEAEKQKAAEATKVAAEAKQ---KAAEATKVAE----- 481
QY 343 GSKQOQSSIGSIRVSMMLDDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARA 402
Db 482 AEKQKAAEATKVAAEAEKQKAAEATK-----EAEKQKAAEATK-----VAEAEKQKAAEATKVAAEAK 526
QY 403 AKAAAGDSDSAAAA---LADAQKALEAAALGKAGQ 431
Db 527 QKAAEATKVAAEAEKQKAAEATKVAAEAEKQKAGE 559

RESULT 15
US-09-115-746-2
; Sequence 2, Application US/09115746
; Patent No. 6228601
; GENERAL INFORMATION:
; APPLICANT: Kirchhoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/115,746
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,894
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-115-746-2

Query Match 6.5%; Score 142.5; DB 3; Length 564;
Best Local Similarity 21.4%; Pred. No. 0.0095;
Matches 97; Conservative 55; Mismatches 164; Indels 137; Gaps 17;

QY 7 PGPIDETERTPPADLSAQGLEASAANKSABAOIAGAEAKPKESKTDVSVERWSILRSVAVN 66
Db 216 PPKSDLVPRGSPQL--QQAENNIITNSKKEMTKLREKVKAEKELDAINRAKLEERN 273
QY 67 ALMSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTDDYKTOAQATYDTFTSTS 126
Db 274 QAYKAAHK---ABEEKAKTFQRLITFSEINILKRP----- 307
QY 127 LADIQAAALVSLQDAVTN-----IKDTATDE-----ETAIAAEWETKNADAVKGAQI 174
Db 308 -----NDKVSNRDKKKNSETAKTDEVEKQRAAEAAKAVETEKORA-----AEA 350

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 05:59:21 ; Search time 105.628 Seconds
(without alignments)
1108.119 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204

Sequence: 1 MWNPIGGPIDEPTPPAD.....QGILNALQIAAAVVSAGV 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pap.*
2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubaa/US09_PUBCOMB.pap.*
10: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pap.*
12: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pap.*
13: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pap.*
14: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pap.*
15: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pap.*
16: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pap.*
17: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pap.*
18: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pap.*
19: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2204	100.0	452	10	US-09-564-479-6
2	2204	100.0	490	10	US-09-564-479-8
3	2204	100.0	651	9	US-09-841-132-405
4	2204	100.0	651	10	US-09-564-479-2
5	2204	100.0	651	10	US-09-841-260-78
6	2204	100.0	651	13	US-10-007-693-78
7	2204	100.0	651	15	US-10-312-273-13
8	2182	99.0	478	15	US-10-289-762-776
9	1866	84.7	583	10	US-09-564-479-4
10	461.5	20.9	647	9	US-09-841-132-436
11	453	20.6	660	10	US-09-841-260-139
12	453	20.6	660	13	US-10-007-693-139
13	231.5	10.5	361	9	US-09-841-132-239
14	219	9.9	1463	9	US-09-971-536-69
15	190	8.6	2478	9	US-09-815-242-5816

16	190	8.6	2478	9	US-09-815-242-12967	Sequence 12967, A
17	189.5	8.6	1831	12	US-10-282-122A-71033	Sequence 71033, A
18	188	8.5	603	12	US-10-282-122A-72004	Sequence 72004, A
19	186.5	8.5	790	12	US-10-282-122A-75142	Sequence 75142, A
20	184	8.3	1122	12	US-10-282-122A-42783	Sequence 42783, A
21	181	8.2	2481	12	US-10-282-122A-43762	Sequence 43762, A
22	178	8.1	1178	14	US-10-128-714-8240	Sequence 8240, A
23	176	8.0	1924	12	US-10-282-122A-43924	Sequence 43924, A
24	173	7.8	2271	12	US-10-282-122A-43946	Sequence 43946, A
25	172.5	7.8	812	12	US-10-282-122A-75242	Sequence 75242, A
26	172	7.8	2344	9	US-09-815-242-12713	Sequence 12713, A
27	171.5	7.8	1879	9	US-09-971-536-70	Sequence 70, Appl
28	170.5	7.7	2368	9	US-09-815-242-5635	Sequence 5635, Ap
29	170.5	7.7	2368	9	US-09-815-242-12389	Sequence 12389, A
30	168.5	7.6	2398	12	US-10-282-122A-70176	Sequence 70176, A
31	168.5	7.6	6281	9	US-09-815-242-12996	Sequence 12996, A
32	168	7.6	993	12	US-10-282-122A-43875	Sequence 43875, A
33	168	7.6	2283	14	US-10-172-502-4	Sequence 4, Appli
34	166.5	7.6	469	12	US-10-282-122A-70453	Sequence 70453, A
35	166.5	7.6	1665	12	US-10-282-122A-71690	Sequence 71690, A
36	166	7.5	2086	9	US-09-815-242-5639	Sequence 5639, Ap
37	166	7.5	5795	9	US-09-815-242-12610	Sequence 12610, A
38	165.5	7.5	385	12	US-10-282-122A-71885	Sequence 71885, A
39	165.5	7.5	454	12	US-10-424-599-232134	Sequence 232134, A
40	164	7.4	6713	12	US-10-282-122A-43811	Sequence 43811, A
41	163.5	7.4	1123	12	US-10-282-122A-70581	Sequence 70581, A
42	163	7.4	1348	12	US-10-282-122A-56877	Sequence 56877, A
43	162.5	7.4	2437	9	US-09-815-242-5834	Sequence 5834, Ap
44	162	7.4	722	9	US-09-815-242-12888	Sequence 12888, A
45	162	7.4	991	9	US-09-815-242-5803	Sequence 5803, Ap

ALIGNMENTS

RESULT 1

US-09-564-479-6
; Sequence 6, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-6

Query Match 100.0%; Score 2204; DB 10; Length 452;
Best Local Similarity 100.0%; Pred. No. 2.9e+160;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MWNPIGGPIDEPTPPADLSAQQEASANKSAEORAGAEAKPKSKTDSVERWSI	60
Db	1	MWNPIGGPIDEPTPPADLSAQQEASANKSAEORAGAEAKPKSKTDSVERWSI	60
QY	61	LRSAVNALMSLADKLGIASSNSSSSSRGADVDSITATPTPPPTFDYKTKQAQYAYDT	120
Db	61	LRSAVNALMSLADKLGIASSNSSSSSRGADVDSITATPTPPPTFDYKTKQAQYAYDT	120

121 IFSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
121 IFSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNDPVVPGKTPAIAQSLVD 240
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNDPVVPGKTPAIAQSLVD 240
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
361 DDAENETASILMSGFRQMIHMENTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
361 DDAENETASILMSGFRQMIHMENTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 2

US-09-564-479-8
; Sequence 8, Application US/09564479
; Publication No. US2003009573A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COHEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: US/09564479
; CURRENT APPLICATION NUMBER: US/09564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR FILING DATE: 2000-05-03
; PRIOR FILING DATE: 1999-05-03
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 8
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-8

Query Match 100.0%; Score 2204; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.2e-160;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEQAQIAGAEAKPKESKTDTSVERWSI 60
DB 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEQAQIAGAEAKPKESKTDTSVERWSI 60
QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFFDDYKTAQATAYDT 120
DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFFDDYKTAQATAYDT 120
QY 121 IFSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
121 IFSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNDPVVPGKTPAIAQSLVD 240
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNDPVVPGKTPAIAQSLVD 240
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300

241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
361 DDAENETASILMSGFRQMIHMENTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
361 DDAENETASILMSGFRQMIHMENTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 3

US-09-841-132-405
; Sequence 405, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: DIAGNOSIS OF CHLAMYDIAL INFECTION
; CURRENT APPLICATION NUMBER: US/09841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 405
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-405

Query Match 100.0%; Score 2204; DB 9; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.8e-160;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEQAQIAGAEAKPKESKTDTSVERWSI 60
DB 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEQAQIAGAEAKPKESKTDTSVERWSI 60
QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFFDDYKTAQATAYDT 120
DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFFDDYKTAQATAYDT 120
QY 121 IFSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
121 IFSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNDPVVPGKTPAIAQSLVD 240
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNDPVVPGKTPAIAQSLVD 240
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSIRVSMML 360
361 DDAENETASILMSGFRQMIHMENTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
361 DDAENETASILMSGFRQMIHMENTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 4

US-09-564-479-2
 ; Sequence 2, Application US/09564479
 ; Publication No. US20030095973A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: COMEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; CURRENT FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/132,270
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR APPLICATION NUMBER: 60/141,276
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 651
 ; TYPE: PR1
 ; ORGANISM: Chlamydia pneumoniae
 ; ORGANISM: Chlamydia pneumoniae

Query Match 100.0%; Score 2204; DB 10; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.8e-160;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAAKSAQAQRIAGAEAKPKESKTSVERWSI 60
 DB 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAAKSAQAQRIAGAEAKPKESKTSVERWSI 60
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDDYKTAQTAYDT 120
 DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDDYKTAQTAYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVTNICKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTNICKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 QY 181 ASDNQAILDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
 DB 181 ASDNQAILDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 DB 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
 DB 361 DDAENETASILMSGFRQMIHMFNTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
 QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 5

US-09-841-260-78
 ; Sequence 78, Application US/09841260
 ; Publication No. US20030175700A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; APPLICANT: Stromberg, Erika Jean
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
 ; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.515
 ; CURRENT APPLICATION NUMBER: US/09/841,260
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 140
 ; SEQ ID NO 78
 ; LENGTH: 651
 ; TYPE: PR1
 ; ORGANISM: Chlamydia pneumoniae
 ; ORGANISM: Chlamydia pneumoniae

Query Match 100.0%; Score 2204; DB 10; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.8e-160;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAAKSAQAQRIAGAEAKPKESKTSVERWSI 60
 DB 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAAKSAQAQRIAGAEAKPKESKTSVERWSI 60
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDDYKTAQTAYDT 120
 DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDDYKTAQTAYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVTNICKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTNICKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 QY 181 ASDNQAILDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
 DB 181 ASDNQAILDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 DB 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
 DB 361 DDAENETASILMSGFRQMIHMFNTENPDQAAQELAAQARAARAAAGDSDSAAALADAQK 420
 QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 6

US-10-007-693-78
 ; Sequence 78, Application US/10007693
 ; Publication No. US20020146776A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
 ; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.515C2
 ; CURRENT APPLICATION NUMBER: US/10/007,693
 ; CURRENT FILING DATE: 2001-12-05
 ; NUMBER OF SEQ ID NOS: 157
 ; SEQ ID NO 78
 ; LENGTH: 651
 ; TYPE: PR1
 ; ORGANISM: Chlamydia pneumoniae
 ; ORGANISM: Chlamydia pneumoniae

Query Match 100.0%; Score 2204; DB 13; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.8e-160;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAAKSAQAQRIAGAEAKPKESKTSVERWSI 60
 DB 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAAKSAQAQRIAGAEAKPKESKTSVERWSI 60

QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPPPTFDDYKTAQAYDT 120
DB 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPPPTFDDYKTAQAYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 180
QY 181 ASDNQAILDSLGKLTSTFLLQALLOSQVANNKKAELLKEMQDNVVPVPGKTPALQAQSLVD 240
DB 181 ASDNQAILDSLGKLTSTFLLQALLOSQVANNKKAELLKEMQDNVVPVPGKTPALQAQSLVD 240
QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAQ 300
DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAQ 300
QY 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSGKQOGSSIGSIRVSMML 360
DB 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSGKQOGSSIGSIRVSMML 360
QY 361 DDAENETASILMSGFRQMHMFNTENPDQAAQELAAQARAAGDSDSAAALADAOX 420
DB 361 DDAENETASILMSGFRQMHMFNTENPDQAAQELAAQARAAGDSDSAAALADAOX 420
QY 421 ALAALGKAGQOQOGLNALGQIASAAVVSAGV 452
DB 421 ALAALGKAGQOQOGLNALGQIASAAVVSAGV 452

RESULT 7
US-10-312-273-13
; Sequence 13, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035W0
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 13
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae

Query Match 100.0%; Score 2204; DB 15; Length 651;
Best Local Similarity 100.0%; Pred. No. 4,8e-160;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
DB 1 MVNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPPPTFDDYKTAQAYDT 120

DB 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPPPTFDDYKTAQAYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 180
QY 181 ASDNQAILDSLGKLTSTFLLQALLOSQVANNKKAELLKEMQDNVVPVPGKTPALQAQSLVD 240
DB 181 ASDNQAILDSLGKLTSTFLLQALLOSQVANNKKAELLKEMQDNVVPVPGKTPALQAQSLVD 240
QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAQ 300
DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAQ 300
QY 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSGKQOGSSIGSIRVSMML 360
DB 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSGKQOGSSIGSIRVSMML 360
QY 361 DDAENETASILMSGFRQMHMFNTENPDQAAQELAAQARAAGDSDSAAALADAOX 420
DB 361 DDAENETASILMSGFRQMHMFNTENPDQAAQELAAQARAAGDSDSAAALADAOX 420
QY 421 ALAALGKAGQOQOGLNALGQIASAAVVSAGV 452
DB 421 ALAALGKAGQOQOGLNALGQIASAAVVSAGV 452

RESULT 8
US-10-289-762-776
; Sequence 776, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment thereof and uses thereof, in particular for the diagnosis, prevention of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 97/20-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 776
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-776

Query Match 99.0%; Score 2182; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.5e-158;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
DB 3 MVNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 62
QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPPPTFDDYKTAQAYDT 120
DB 63 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPPPTFDDYKTAQAYDT 122
QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 180
DB 123 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGAIITELAKY 182
QY 181 ASDNQAILDSLGKLTSTFLLQALLOSQVANNKKAELLKEMQDNVVPVPGKTPALQAQSLVD 240
DB 183 ASDNQAILDSLGKLTSTFLLQALLOSQVANNKKAELLKEMQDNVVPVPGKTPALQAQSLVD 242
QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAQ 300
DB 243 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAQ 302
QY 301 KKFPPDSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSGKQOGSSIGSIRVSMML 360

Db 303 KKFDPSPILQAEQWVIAEKDKLVKIPADGSDVNPFTTGGSKQGGSSIGTSIRVSMLL 362
 Qy 361 DDAENETASILMSGFQWIMHFNTEPDSQAQOELAAQAPAAKAGDSDSAAALADAQK 420
 Db 363 DDAENETASILMSGFQWIMHFNTEPDSQAQOELAAQAPAAKAGDSDSAAALADAQK 422
 Qy 421 ALEAALGKAGQOQGIILNALGQIASAAV 447
 Db 423 ALEAALGKAGQOQGIILNALGQIASAAV 449

RESULT 9
 US-09-564-479-4
 ; Sequence 4, Application US/09564479
 ; Publication No. US20030095973A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: WOMEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; PRIOR FILING DATE: 2000-05-03
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR FILING DATE: 60/141,276
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 583
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-564-479-4

Query Match 84.7%; Score 1866; DB 10; Length 583;
 Best Local Similarity 100.0%; Pred. No. 3e-134;
 Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 69 MSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFFDYKTOAQATYDTTFTSLA 128
 Db 1 MSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFFDYKTOAQATYDTTFTSLA 60
 Qy 129 DQAAVLSQDAVNIKDTAATDETAIAEWETKNADAVKVGQAQITELAKYASDNQAIL 188
 Db 61 DQAAVLSQDAVNIKDTAATDETAIAEWETKNADAVKVGQAQITELAKYASDNQAIL 120
 Qy 189 DSIQKLTSLDLOALQSVANNKKAELLKEMQDNVPPVPGKTPAIAQSLVDQDTATQ 248
 Db 121 DSIQKLTSLDLOALQSVANNKKAELLKEMQDNVPPVPGKTPAIAQSLVDQDTATQ 180
 Qy 249 IEKDGNAIRDAYFAGQNASGAVENAKNSNISNDSKAAIATAKTQIAEAKKFPDPSI 308
 Db 181 IEKDGNAIRDAYFAGQNASGAVENAKNSNISNDSKAAIATAKTQIAEAKKFPDPSI 240
 Qy 309 LQEAQWVIAEKDKLVKIPADGSDVNPFTTGGSKQGGSSIGTSIRVSMLLDDAENETA 368
 Db 241 LQEAQWVIAEKDKLVKIPADGSDVNPFTTGGSKQGGSSIGTSIRVSMLLDDAENETA 300
 Qy 369 SILMSGFQWIMHFNTEPDSQAQOELAAQAPAAKAGDSDSAAALADAQKALEALGK 428
 Db 301 SILMSGFQWIMHFNTEPDSQAQOELAAQAPAAKAGDSDSAAALADAQKALEALGK 360
 Qy 429 AGQOQGIILNALGQIASAAV 452
 Db 361 AGQOQGIILNALGQIASAAV 384

RESULT 10
 US-09-841-132-436
 ; Sequence 436, Application US/09841132

; Patent No. US20020061848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.469C8
 ; CURRENT APPLICATION NUMBER: US/09/841,132
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 599
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 436
 ; LENGTH: 647
 ; TYPE: PRT
 ; ORGANISM: Chlamydia trachomatis serovar D
 US-09-841-132-436

Query Match 20.9%; Score 461.5; DB 9; Length 647;
 Best Local Similarity 29.7%; Pred. No. 1e-26;
 Matches 138; Conservative 86; Mismatches 192; Indels 49; Gaps 13;
 Qy 2 VNPICGPIDETERTPPADLSAQGLEASAANKSAEQRIAGAEAKPKESKTDVSVERVSL 61
 Db 15 MNPIINGQI-----ANSETKESKSEA-----SPSASSSVSSWSFL 52
 Qy 62 RSVAVALMSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFFDYKTOAQATYDT 120
 Db 53 SSAKHALISLRD--AILKNKSSPTDSLQLEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110
 Qy 121 IFTSTSLADIAAALVSLQDAVNIKDTAATDEB-TAIAEWETKNADAVKVGQAQITELAK 179
 Db 111 LENATTLAEYETKMDLMAALQDMERLAKQAEVTRIKALQEKQ-----EVIDKLNQLVK 166
 Qy 180 YASDNQAILDSLGLKLTSLDLOALQSVANNKKAELLKEMQDNVPPVPGKTPAIAQSLV 239
 Db 167 LEKQNTKLTETLTTSADQIPAINSOLEINKNSADQIIXOLEGQNI---SYEAVLTNAG 223
 Qy 240 DOTDATATQIEKDGNAIRDAYFAGQNASGAVENAKNSNISNDSKAAIATAKTQIAEA 299
 Db 224 EVIKASSEAGIKLQALQSVLDAGDQSAAVLQAOQNNSPDNIAATKCLIDAAETKVNEL 283
 Qy 300 QKK---FPDPSILQEAQWVIAEKDKLVKIPADGSDVNPFTTGGSKQGGSSIGSI-- 354
 Db 284 KQEHGTGLTDSPLVKKABEQISQAKQIQEIKPS-GSDIPIVGPS-GSAAGSAGSAGALKS 341
 Qy 355 -----RVSMLLDDAENETASILMSGFQWIMHFNTEPDSQAQOELAAQAPAAK---AA 406
 Db 342 SNNSGRISLILDDVDNEMAIAMQGFMSMLEQFNVPNPATAKELQAMEAQITAMSDQV 401
 Qy 407 GDSAAALADAQKALEALGKAGQOQGIILNALGQIASAAVVSAG 451
 Db 402 ADGELPAEIOAIKDALAQAL-KQPSDTGLATAMQGVAFAAAKVGG 445

RESULT 11
 US-09-841-260-139
 ; Sequence 139, Application US/09841260
 ; Publication No. US20030175700A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; APPLICANT: Stromberg, Erika Jean
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND DIAGNOSIS
 ; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.515
 ; CURRENT APPLICATION NUMBER: US/09/841,260
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 140
 ; SEQ ID NO 139
 ; LENGTH: 660
 ; TYPE: PRT
 ; ORGANISM: Chlamydia trachomatis

US-09-841-260-139

Query Match 20.6%; Score 453; DB 10; Length 660;

Best Local Similarity 28.9%; Pred. No. 4.6e-26;

Matches 136; Conservative 94; Mismatches 193; Indels 48; Gaps 14;

QY 6 GPGPIDETRTTPADLSAAGLEASAAKSAEABQRTAGAEAKPKESKTDVSERKSIIRSAV 65

Db 11 GPESVSSNQSNMNPINGQIASNETKSTKA-----SEASP--SASSSVSSWSFLSSAK 63

QY 66 NALMSADKLGIIASSNSSSTSRK-ADVDTTATAPTTPPTFDDYKTKQAQYATDTFTS 124

Db 64 NALISLRD--AILKNSSPTDSLSQLEASTSISTVRVAAKYDEAKSNFTAKSGLENA 121

QY 125 TSLADIQAALVSLQDVTNIKDTAATD-----EET-----ATAEWETKNADAVKVGQA 173

Db 122 KTLAEYETKMADLMAALQDWERLANSDPNNHTEEVNNIKKALEAQKDTID-----K 173

QY 174 ITELAYASDNQAILDSLGKLTSPDLLQAALLQSVANNKKAELLKEMQDNPVVPGKTPA 233

Db 174 LNKLVTLQNKQKSLTEVLTKTDSADQIPAINSQLINKNSADQIIKDLERQNI---SYEA 230

QY 234 IAQSLVDQTDATATQIEKQGNARDAVFAGQNASGAVENAKSNNSISNIDSAKAAIATAK 293

Db 231 VLTNAGEVTKASSEAGIKLQALQSIVDAGDQSOAAVLQAQNNSPDNIAATKELLIDAE 290

QY 294 TQIAEAQKK---FPDSPILQAEQMVIOAEKDLKNIKPADGSDVNPNGTTVGGSKQCGSS 350

Db 291 TKVNELKQHTGLTDSPLVKABEQISOAQKDIQIKPS-GSDIPIVGPS-GSAASAGSA 348

QY 351 IGS1-----RVSMILDDAENETASILMSGFQMIHNTENPDPSQAQOELAAQARAA 403

Db 349 AGALKSSNNSGRISLLDDVDNEMAALQGFRRSMIEQFNVNPNPATAKELQAMEAQLTAM 408

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 174 ITELAYASDNQAILDSLGKLTSPDLLQAALLQSVANNKKAELLKEMQDNPVVPGKTPA 233

Db 174 LNKLVTLQNKQKSLTEVLTKTDSADQIPAINSQLINKNSADQIIKDLERQNI---SYEA 230

QY 234 IAQSLVDQTDATATQIEKQGNARDAVFAGQNASGAVENAKSNNSISNIDSAKAAIATAK 293

Db 231 VLTNAGEVTKASSEAGIKLQALQSIVDAGDQSOAAVLQAQNNSPDNIAATKELLIDAE 290

QY 294 TQIAEAQKK---FPDSPILQAEQMVIOAEKDLKNIKPADGSDVNPNGTTVGGSKQCGSS 350

Db 291 TKVNELKQHTGLTDSPLVKABEQISOAQKDIQIKPS-GSDIPIVGPS-GSAASAGSA 348

QY 351 IGS1-----RVSMILDDAENETASILMSGFQMIHNTENPDPSQAQOELAAQARAA 403

Db 349 AGALKSSNNSGRISLLDDVDNEMAALQGFRRSMIEQFNVNPNPATAKELQAMEAQLTAM 408

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQGGILNALGQIASAAVVSAG 451

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

QY 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

Db 409 SDQLVGADGELPAEIQAIKDALAQAL-KQPSADGLATANGQVAFPAAKVGG 458

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:39:25 ; Search time 8.31079 Seconds
(without alignments)
2831.945 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204

Sequence: 1 MNPDPGPDIDERTPPAD.....QGILNALGQIASAAVVSAGV 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	184	8.3	1120	1	STFR_ECOLI
2	176.5	8.0	545	1	HTR5_HALN1
3	160	7.3	535	1	HTR1_HALN1
4	156	7.1	810	1	HTR4_HALN1
5	156	7.1	810	1	HTR4_HALN1
6	154.5	7.0	1238	1	SRCC_RHOCA
7	151	6.9	641	1	HRPK_PSESY
8	149	6.8	778	1	HTR6_HALSA
9	148.5	6.7	763	1	HTR2_HALN1
10	148.5	6.7	764	1	HTR2_HALSA
11	147.5	6.7	1528	1	SPAA_STRDO
12	147	6.7	505	1	FLJB_SALTY
13	146	6.6	778	1	HTR6_HALN1
14	145	6.6	1565	1	PAC_STRMU
15	145	6.6	1714	1	SYEP_DROME
16	144	6.5	881	1	PRY3_YEAST
17	144	6.5	1562	1	SPAP_STRMU
18	143.5	6.5	758	1	INMT_HUMAN
19	141.5	6.4	2542	1	TLN2_HUMAN
20	141	6.4	535	1	HTR1_HALSA
21	141	6.4	819	1	NRFC_HAEIN
22	139.5	6.3	774	1	STFL_LAMED
23	139	6.3	705	1	CMB4_BACSU
24	139	6.3	1637	1	MRSP_STRAU
25	138.5	6.3	1969	1	MYSA_CABEL
26	138	6.3	1609	1	LMG1_HUMAN
27	136	6.2	1607	1	LMG1_MOUSE
28	135.5	6.1	1061	1	TRC4_ECOLI
29	135	6.1	1306	1	MSB2_YEAST
30	134	6.1	376	1	FLAC_VIBAN
31	134	6.1	490	1	CE05_ECOLI
32	134	6.1	564	1	M12_STRPY
33	134	6.1	8545	1	ANCI_CABEL

34 133.5 6.1 1013 1 A60D_DROME
35 133.5 6.1 1411 1 TCOF_HUMAN
36 133 6.0 500 1 FLJB_SALAE
37 133 6.0 797 1 VGLX_HSVBB
38 133 6.0 1205 1 SMC_MYCTU
39 132.5 6.0 2492 1 TALA_DICDI
40 132.5 6.0 2541 1 TLNI_MOUSE
41 132 6.0 475 1 MOTD_RHIME
42 132 6.0 757 1 IMMT_MOUSE
43 132 6.0 2843 1 APC_HUMAN
44 132 6.0 2845 1 APC_MOUSE
45 131.5 6.0 573 1 YEFY_YEAST

ALIGNMENTS

RESULT 1
STFR_ECOLI
ID STFR_ECOLI STANDARD; PRT; 1120 AA.
AC P76072; P77560;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Side tail fiber protein homolog from lambdaoid prophage Rac.
GN STFR OR H1372.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=97425617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
[2]
RC SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map";
RL DNA Res. 3:363-377(1996).
CC -!- SIMILARITY: Belongs to the tail fiber family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000234; AAC74454.1; ALT_INIT.
CC EMBL; D90774; BAA14966.1; -.
CC EMBL; D90775; BAA14975.1; -.
CC PIR; G64887; G64887.
CC EcoGene; EG13370; stfr.
CC InterPro; IPR008969; CarboxypepD_reg.
CC InterPro; IPR005003; Phage_fiber.
CC InterPro; IPR005068; Phage_fiber_2.
CC Pfam; PF03335; Phage_fiber; 6_2.
CC Pfam; PF03406; Phage_fiber_2; 1.

```
KW Hypothetical protein; Fiber protein; Repeat; Complete proteome.
SQ SEQUENCE 1120 AA; 113779 MW; 542E59D71EE795B4 CRC64;

Query Match
Best Local Similarity 8.3%; Score 184; DB 1; Length 1120;
Matches 102; Conservative 69; Mismatches 201; Indels 78; Gaps 14;

QY 8 GPIDEITERPPA---DLQAQGLEASAAANKSAQAQRIAGAEAKPKESKTSVERWILRS 63
Db 91 GAWTEDDARPEALRRLELM---VEEVARNASAVAQNTAAKKASDASTISAREAAH 147
QY 64 AVNALMSLADKGLIASNSSSSTRSADVDSTATPTPPPTFFDYKQAOYDTIT 123
Db 148 AADSARAASTAGQNAS-SQAASSAGTASTKATASKASAAAESEKSAATAGAAKT 206
QY 124 S--TSLADIQAALVSLQDAVTNKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKYA 181
Db 207 SETNASASLSQATSTASTATTKASE-AATSARDAASKEAKKSE-----TNASSA 257
QY 182 SDNQAILDSGLKLTSPDLLQAALLOSANNKAAELKEMQDNVPVVGKTPAIAQSLVDQ 241
Db 258 SSAASATAAGN-----SAAKATSETVARSSETAAGQSAAGSATAAASSASAAS 310
QY 242 TDATATQIEKDGNAIRDAPAGQNASGAVENAKSNNSISNIDSAKAAIATKQIAEAK 301
Db 311 TSA-----GOASASATAAGKSA-----ESAASSASTATTKAGEATE 346
QY 302 KFPDPSILOAEQWVIQAEKDLKNIPADGSDVPNPGTTVGGSKQQGSSIGSRVSMLLD 361
Db 347 Q-----ASAAASASAAKTSSETNAKASETSABESSKTAASASSASAASAK--D 398
QY 362 DAENETASILMSFROMHMFNTENPDQAAQELAAQA-----RAAKAAGDSDSAAAL 415
Db 399 EATROASAASSATTAATKATKAAGSATAAQAQSKTAESAAATRAETAAKRAEDIAVA 458
QY 416 ADAQKALEALGKAGQOGLNALGQIASA 445
Db 459 EDASTT-----KKGIV-----QLSSA 474

RESULT 2
HTR5 HALN1 STANDARD; PRT; 545 AA.
AC Q48318; Q9HP85;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Halobacterial transducer protein V.
GN HTR7 OR HTPV OR VNG1.759G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091, 2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1;
RX MEDLINE=20504483; PubMed=11016950;
RA NG W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsen V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Allam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=H.salinarium; STRAIN=S9;
RX MEDLINE=96275896; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
```

```
RA Oesterhelt D.;
RT "A family of halobacterial transducer proteins.";
RL FEMS Microbiol. Lett. 139:161-168(1996).
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal transduction.
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 1 HAMP domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005080; AAG19985.1; -.
DR EMBL; X95589; CAAG4842.1; -.
DR PIR; E84327; E84327.
DR PIR; T46811; T46811.
DR HSSP; P02942; 1QU7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR004090; Me_Chemtaxis.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PRO0260; CHEMTRNSDUCR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1. Complete proteome.
KW Transducer; Transmembrane; Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT DOMAIN 164 217 HAMP.
FT DOMAIN 236 476 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 545 AA; 57070 MW; FDB70389C2F428B CRC64;

Query Match
Best Local Similarity 8.0%; Score 176.5; DB 1; Length 545;
Matches 98; Conservative 79; Mismatches 195; Indels 87; Gaps 17;

QY 12 ETERTPPDLQAQLEASAAANKSAQAQRIAGAEAKPKESKTSVERWILRSVAVNALMSL 71
Db 118 EEERA-----EAERAREKAEQKQAEAE---QTAEASAKODARERSAEIEQLAADLESQ 169
QY 72 ADKLGIASSNSSSTSRGADVSTTATPTPPPT-FDDYKQAOYDTI-----FTST 125
Db 170 ATEVG-ATLEASADGDLTARVDATTDAEIAEAVATVNDMLTMTERTIDEIOGFSTNVT 228
QY 126 SLADIQAALVSLQDAVTNIKOT-----AATDBE---TAIAAEWETKNADAVKVGQAQITE 176
Db 229 ASREATAGAKETQDASQTVSESVQEIAGTDQREQLESVABEMDSYSATVEEVAATAQS 288
QY 177 LAKYASDNCAILDSLGKLT-----SFDLLQAALLOSANNKAAELKEMQDNVPVPGK 230
Db 289 VADTAADTTDVA-TAGKQTAEDAIDAIDAVQETWQTTVANVDALEDTLTEIDD----- 340
QY 231 TPAIAQSLVDQDATATQIEKDNRAIRDAPAGQNASG-----AVENAKSNNSI 279
Db 341 IAEILSDTAEOINMLA--LNANIEARAGSGGSGNGDGFVAVVDEVKELATESQSAKDI 398
QY 280 SN-ISAAXAATATKQIAEAKPKFPDPSILQAEQWVIQAEKDLKNTPADGSDVPNPG 338
Db 399 AELIEEVQSQTATVEEI-----RVAEQRVNDGAAAVEETVDAPGAVTENIQ 445
QY 339 TTVGSKQGGSSIGSRVSMLLDDEAENETASILMSGFRQMIFNTENPDSDQAQOELAA 398
Db 446 ETTDGVQE-----ISQAMDEQAQRSRVSS-----VDDIATISQATAD 484
QY 399 QARAAKAGAGDSDSAAALADAAQKALEALGKAGQOGLNAL 437
Db 485 RAENTVSAASEEQ-TASITEVTSLSQSLAAQAQADTLERLN 522
```



```

RESULT 3
HTR1 HALN1
ID HTR1 HALN1 STANDARD; PRT; 535 AA.
AC P33741; Q9HPP6;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sensory rhodopsin I transducer (HTR-I) (Methyl-accepting phototaxis
DE protein I) (MPP-I)
DE
GN HTR1 OR HTR1 OR VNGI1659G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and
OS Halobacterium halobium.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091, 2242;
RN
RN SEQUENCE FROM N.A.
RC STRAIN=NRC-1;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN
RN
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-10; 350-377 AND 457-476.
RC SPECIES=H.halobium; STRAIN=FX5R;
RX MEDLINE=93101637; PubMed=1465418;
RA Yao V.J., Spudich J.L.;
RT "Primary structure of an archaeobacterial transducer, a
RT methyl-accepting protein associated with sensory rhodopsin I.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:11915-11919(1992).
CC
CC -!- FUNCTION: TRANSDUCES SIGNALS FROM THE PHOTOTAXIS RECEPTOR
CC SENSORY RHODOPSIN I (SR-I) TO THE FLAGELLAR MOTOR. RESPONDS TO
CC LIGHT CHANGES THROUGH THE VARIATION OF THE LEVEL OF METHYLATION.
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE005075; AAG19913.1; -
CC EMBL; L05603; AAA72315.1; -
CC PIR; A47190; A47190.
CC PIR; E84318; E84318.
CC HSSP; P02942; 1QU7.
CC InterPro; IPR004089; Chmtaxis_transd.
CC InterPro; IPR003660; HAMP.
CC InterPro; IPR004090; Me_chmotaxis.
CC Pfam; PF00672; HAMP; 1.
CC Pfam; PF00015; MCPsignal; 1.
CC PRINTS; PR00260; CHEMTXNSDUCR.
CC SMART; SM00304; HAMP; 2.
CC SMART; SM00283; NA; 1.
CC PROSITE; PS01011; CHEMOTAXIS_TRANSDUC_2; 1.
CC PROSITE; PS00885; HAMP; 2.
KW Transducer; Photoreceptor; Transmembrane; Methylation;
KW Complete proteome; Repeat.
FT INIT MET 0 0
FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 14 28

```

```

FT DOMAIN 29 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 54 POTENTIAL..
FT DOMAIN 55 535 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 54 106 HAMP 1.
FT DOMAIN 148 201 HAMP 2.
FT DOMAIN 220 458 METHYL-ACCEPTING TRANSDUCER.
FT MOD_RES 265 265 METHYLATION.
FT MOD_RES 272 272 METHYLATION.
FT MOD_RES 279 279 METHYLATION.
FT MOD_RES 463 463 METHYLATION.
FT MOD_RES 472 472 METHYLATION.
SQ SEQUENCE 535 AA; 56544 MW; B9945E4F66A9D091 CRC64;

Query Match 7.3%; Score 160; DB 1; Length 535;
Best Local Similarity 19.1%; Pred. No. 0.15;
Matches 90; Conservative 81; Mismatches 189; Indels 110; Gaps 15;

QY 8 GPIDEETPTPADLSAAGLEASAANKSAE--AORIAGAEEAKPKE---SKTDSVERW--- 58
Db 103 GRUNEMERT-RADLEETQAEATAREAEQAKQAEQAAREARELAATYQDTAKRYGETM 161
QY 59 -----SILRSVNALMSLADKLGIASSSSST 86
Db 162 EAAATGDLTORVDVTDHEAMETVGTAFNOMDDLOATVTVTTVADEIEAKTERMSET- 220
QY 87 SRSDVDVST---TATAPTPPTFDYKQAOATYDTIFTSTSLADIQAALVSLQDAVTN 143
Db 221 --SADIASEAGDVTEAVKIESQANDQRTELDAAAD-----DVQQVSAEAEIAAT 269
QY 144 IKDTATATDEETAIAAE-----WETKNADAVKVGQITELAKYASDNOAIL 188
Db 270 IDLASRSEDVATASDAARDSSKSALEDEMSIETEVDDAVGQVEQLRDQVABITDVIDVI 329
QY 189 DSGKITSFDLLOALLQSVANN-----KAEALLKEMQDNVVPVGTPTAQAQ 237
Db 330 TDIGEQTNLNALNASIEAARAGAGNADGDFSVVADEVKOLAETQDR---ANEIAAVVEK 386
QY 238 LVDQTDATATQIEKGNVNRDAYFAGQNASGAVENAKSNNSINISAKAATAKTQIA 297
Db 387 VTAQTDVNTASIQOTRTV-----ESGETVE-----STLRDRTIADSIASVNSID 434
QY 298 EAQKKEPDSPILOEAEQWIAEKDKAKNIKPADGSDVPNPGTTVGSGSKQGSSTGSRVS 357
Db 435 EIQTSTSEQAETVQSTATSVVERVAGLSDDTTALASDAE---SAVIGQESAEELAA--- 487
QY 358 MLDDAENETASILMSFROMIHMFENPDPSQAQOELAAQARAQAAG 407
Db 488 -SLEQFQNTAVEQLQS-----RVASFVATEDSETAGSGVEQPVNRAGADG 532

RESULT 4
HTR4 HALN1 STANDARD; PRT; 810 AA.
AC Q9HPB4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Halobacterial transducer protein IV.
GN HTR5 OR HTPIV OR VNGI1760G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN
RN SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA

```

```

RA Ebbardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC !- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction (by similarity).
CC !- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC !- SIMILARITY: Contains 2 HAMP domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB005080; AAG19986.1; -.
DR PIR; F84327; F84327.
DR HSSP; P02942; 1QU7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 2.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS00111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 2.
DR Transducer; Transmembrane; Complete proteome; Repeat.
KW DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 POTENTIAL.
FT DOMAIN 60 323 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 324 344 POTENTIAL.
FT DOMAIN 345 810 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 345 397 HAMP 1.
FT DOMAIN 439 493 HAMP 2.
FT DOMAIN 512 748 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 810 AA; 85219 MW; 4BF36E4B7D2BD80 CRC64;

Query Match
Best Local Similarity 7.1%; Score 156; DB 1; Length 810;
Matches 102; Conservative 73; Mismatches 170; Indels 118; Gaps 22;

QY 27 EASAANKSA-EAQRIGAEAKPKESKTSVERWSILRSVAVNLMADKLGIASSNSSSS 85
Db 398 EAFAAEQATEAQDAEAE---RERAEADAREADAKADAEL-----AAELEAQ 444

QY 86 TERSADVSTTATAPPPPTFDYKTAQTAVIDTFTSTSLADIQAALVSLQDAVTNIK 145
Db 445 AERYSDVMAACADGDLTRRMPADDTNEAMAAIAASF-NEMLAQWHEHTIIDIQE---FA 499

QY 146 DTAATDEETAIAAEWETKNAADAVKVGQIT---ELAKYASDNQAILDSL-GKLTSPDLL 200
Db 500 DAVATASEEA-----EVGAADAREASQVSEVQETAGADEORNMLDVTSGEMT--DL- 551

QY 201 QAALQSIVANNKAAELKEMQNPVPGKTPAIAQSLVDQDTATATQIEKGNIAIDAY 260
Db 552 -SAAIEVA-----ASADSVAEHSQTA-EIARDGE----- 580

QY 261 FAGQNSAGVAVENAKSNNSISNIDSAKAAIAITAKTQTAABQKFPD--SPILQAEQWVIQ 318
Db 581 ---QTAEADAIE--RSLTVQEAIDATVQNVBALDDQMAEI-SEIVDGLISDAEQTNMLAN 634

QY 319 AE-----KDLKNIKPADGSDVENPQTVGGSKQGGSSIGSIRVSM 358
Db 635 ANTEARADYSGDGFVAVDEVKDLAEETQESAGDIERRITEV--QSQTATVAEAAAE 692

QY 359 LLDDAENETASILMSGFRQMI-HMFNT-----ENPDSQAQQE----- 395
Db 693 ESMADGIDAVEEVVDAFTAVSDHADETDGTGVEISDITDDQAASTEBAVSMTEVADLSD 752

QY 396 -LAAQARAANKAGDDSAALADAKALEALGKAGQCCQILN 437
Db 753 STAGEAQSVSAABEEQ-AASMSEISDSVESLSGQAEQLKALLS 794

```

```

RESULT 5
HTR4_HALSA
ID HTR4_HALSA STANDARD; PRT; 810 AA.
AC Q48317;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Halobacterial transducer protein IV.
GN HTR5 OR HTRIV.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_taxid=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S9;
RX MEDLINE=96275996; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
RA Oesterhelt D.;
RT "A family of halobacterial transducer proteins.";
RL FEMS Microbiol. Lett. 139:161-168(1996).
CC !- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction.
CC !- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC !- SIMILARITY: Contains 2 HAMP domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95589; CAA64841.1; -.
DR PIR; T46810; T46810.
DR HSSP; P02942; 1QU7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 2.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS00111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 2.
DR Transducer; Transmembrane; Repeat.
KW DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 POTENTIAL.
FT DOMAIN 60 323 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 324 344 POTENTIAL.
FT DOMAIN 345 810 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 345 397 HAMP 1.
FT DOMAIN 439 493 HAMP 2.
FT DOMAIN 512 748 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 810 AA; 85207 MW; 37B0F6046A3D9BA CRC64;

Query Match
Best Local Similarity 7.1%; Score 156; DB 1; Length 810;
Matches 101; Conservative 74; Mismatches 170; Indels 118; Gaps 22;

QY 27 EASAANKSA-EAQRIGAEAKPKESKTSVERWSILRSVAVNLMADKLGIASSNSSSS 85
Db 398 EAFAAEQATEAQDAEAE---RERAEADAREADAKADAEL-----AAELEAQ 444

QY 86 TERSADVSTTATAPPPPTFDYKTAQTAVIDTFTSTSLADIQAALVSLQDAVTNIK 145
Db 445 AERYSDVMAACADGDLTRRMPADDTNEAMAAIAASF-NEMLAQWHEHTIIDIQE---FA 499

QY 146 DTAATDEETAIAAEWETKNAADAVKVGQIT---ELAKYASDNQAILDSL-GKLTSPDLL 200
Db 500 DAVATASEEA-----EVGAADAREASQVSEVQETAGADEORNMLDVTSGEMT--DL- 551

```

```

QY 201 QALLOSVAANNKAAELLKEMQDNVVPVGTGTPATAQSLVPOQTATATATQIIEKQGNIRDAY 260
D 552 -SAAIEBVA-----AGADSVABSHQTA-EIARDGE----- 580
QY 261 FAGQASGAVENAKSNNSISNIDSAKAAIATPAKQIAEAQKFPD--SPILQEAQEQVVIQ 318
D 581 ---QTAEDALERSUSVQEA--IDATQVNEALDDQMAEI-SEIVDLISDIAEQNMMLALN 634
QY 319 AE-----KDLKNIKPADGSDVPNPGTIVGGSKQQGSSISGIRVSM 358
D 635 ANIEAARADKSGDGFVAVVADEVKDLAEBETQBSAGDIERRITEV--QSQTTATVAEAAAE 692
QY 359 LLDABNETASILMSGFRQMI-HMFNT-----ENPDSQAAQOE----- 395
D 693 ESM DAGIDAVEEVYDAFTAVSDHADETDGTGVQEISDTTDDQASSTFEAVSMTEEVADLSD 752
QY 396 -LAAQARAANKAGDSDSAAALADAAQKALEALGKAGQOQGLN 437
D 753 STAGEAQSVAABEQ-AASMSSEISDSVESLGGQEQKALLS 794

RESULT 6
SBCC RHOC
ID SBCC RHOC STANDARD; PRT; 1238 AA.
AC 088032; 2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Nuclease sbcCD subunit C.
GN SBCC.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OX Rhodobacteraceae; Rhodobacter.
RN NCBI_TaxID=1061;
[1]
SEQUENCE FROM N.A.
RC STRAIN=SB1003 / St. Louis;
RX MEDLINE=97404404; PubMed=9256491;
RA Vicek C., Faces V., Maltsev N., Faces J., Haselkorn R., Fonstein M.;
RT "Sequence of a 189-kb segment of the chromosome of Rhodobacter
RL capsulatus SB1003."
RC Proc. Natl. Acad. Sci. U.S.A. 94:9384-9388 (1997).
CC -!- FUNCTION: SbcCD cleaves DNA hairpin structures. These structures
CC can inhibit DNA replication and are intermediates in certain DNA
CC recombination reactions. The complex acts as a 3'->5' double
CC strand exonuclease that can open hairpins. It also has a 5'
CC single-strand endonuclease activity (By similarity).
CC -!- SUBUNIT: Heterodimer of sbcC and sbcD (By similarity).
CC -!- SIMILARITY: Belongs to the SMC family. SbcC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF010496; AAC16118.1; -.
CC PIR; T03465; T03465.
CC DR InterPro; IPR003439; ABC transporter.
CC KW Hydrolase; Nuclease; Exonuclease; Endonuclease; DNA replication;
CC DNA recombination; ATP-binding; Coiled coil.
CC FT NP_BIND 37 44
CC FT DOMAIN 395 438 COILED COIL (POTENTIAL).
CC FT DOMAIN 466 487 COILED COIL (POTENTIAL).
CC FT DOMAIN 521 600 COILED COIL (POTENTIAL).
CC FT DOMAIN 724 770 COILED COIL (POTENTIAL).
CC FT DOMAIN 901 943 COILED COIL (POTENTIAL).
CC FT DOMAIN 1019 1052 COILED COIL (POTENTIAL).
CC SEQUENCE 1238 AA; 128046 MW; 2B6BC6C63A85AE1 CRC64;

```

```

Query Match 7.0%; Score 154.5; DB 1; Length 1238;
Best Local Similarity 22.1%; Pred. No. 0.82;
Matches 126; Conservative 79; Mismatches 223; Indels 143; Gaps 23;
QY 6 GP--GPTDETERTPP-----ADLSAQGLEASAAANKSAEAOQRI---AGAEAK 46
D 293 GPARAQDLRLERAPLHLPHQAAAAEAERLDAARSGNAASASRLAAQALATFADAAAR 352
QY 47 PKESKTDSEVERNSILRSVAVNALMSLADKLGIASSNSSSTSRSD-----VDSTT 96
D 353 ATAAGTETEEAFKAFGRINDRAAALDAQIATAATEAARTRAAETARAAGIRRAEADL 412
QY 97 ATAPT-----PPPTFDYDKTCAQTAYDTITFTSLADIQAALVSLQDAVTNKTAA 149
D 413 AAATRAQAQAFQAARLAELAAQAPLADDPQLRRDLADHRAACAAQAAAAATAAQ--AA 470
QY 150 TDEETAAAEWETKNADAVKVGQITELAKYASDNOAILDSL-----GKLTSPDLL 200
D 471 QDRAQALAQEAAAAAARDAADQADSNALAKAAALQALAPLEAAHPPARTADLARIETD 530
QY 201 QAAALLOSVAANNKAAELLKEMQDNVVPVGTGTPATAQ-----SLVDQTDATATQI- 249
D 531 LAELTRALRDCAEAAALGAAQRAATAAATAAALARAQESAQAQDLDRATQIAALTAPL 590
QY 250 EKDGNAIRDA-----YFAGQ-----NAGAVEN 272
D 591 EQADLALSDAARSURAQLSAGSPCVGGALEHPTPABAGLAHLAERLURADQAAARGAAQA 650
QY 273 AKSNNSISNIDSAAKAAIATATAKTAIEAQAQKFPDPSPIQEAQEQVIOAEKDLKNIKPA-DG 331
D 651 AR--DALTAAGGARATAEARGTQAEEDQRAQTAEEAARAANGDTQPRVSARPLAPALPG 708
QY 332 SDVNPGETTVGGSK-----QQSSSIGSRVSM---LDDAENE-----TASILMSGFRQ 377
D 709 T--PDP--TALAAAQDRLCALQTAAEAQAABISALRTLTEARDRELRALLAHRTRE 765
QY 378 MIHMENTENPDSQAQQLAAQARAQAAG-----DSDAAAALADAQ 419
D 766 RLAV----QQAETAQEAALAEARTEAARAAARRDGLALALAPALARAGEDDPAAPGLAERL 820
QY 420 KALEAALGKA--GQQ--QGILNALG-QIASA 445
D 821 AATVSAVGARTGLQAQAEALSALAPQLANA 851

RESULT 7
HRPK PSEY
ID HRPK PSEY STANDARD; PRT; 641 AA.
AC P41501;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Pathogenicity locus protein hrpK.
GN HRPK.
OS Pseudomonas syringae (pv. syringae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=321;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pss61;
RX MEDLINE=94148760; PubMed=8106313;
RT Xiao Y., Heu S., Yi J., Lu Y., Hutcheson S.W.;
RT "Identification of a putative alternate sigma factor and
RT characterization of a multicomponent regulatory cascade controlling
RT the expression of Pseudomonas syringae pv. syringae Pss61 hrp and
RT hma genes."
RL J. Bacteriol. 176:1025-1036 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pss61;
RX MEDLINE=94100578; PubMed=8274770;
RA Heu S., Hutcheson S.W.;

```



```

RESULT 9
ID HTR2 HALNI STANDARD; PRT; 763 AA.
AC Q9HPB1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis
DE protein II) (MPP-II).
GN HTR2 OR VNG1765G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welte R., Goo Y.A.,
RA Lethausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Eberhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,
RA "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: Transduces signals from the phototaxis receptor sensory
CC rhodopsin II (SR-II) to the flagellar motor. Responds to light
CC changes through the variation of the level of methylation. Also
CC acts as a chemotransducer (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005080; AAC19989.1; -.
DR F1R; A84328; A84328.
DR HSP; F02942; LQUT.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCFsignal; 1.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; NA; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 2.
DR Transducer; Photoreceptor; Chemotaxis; Transmembrane; Methylation;
KW Complete proteome; Repeat.
FT INIT_MET 0 0 BY SIMILARITY.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 36 POTENTIAL.
FT DOMAIN 37 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 297 POTENTIAL.
FT DOMAIN 298 763 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 302 354 HAMP 1.
FT DOMAIN 396 449 HAMP 2.
FT DOMAIN 468 704 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 763 AA; 78911 NW; CF7A8FF04DF309A CRC64;
Query Match 6.7%; Score 148.5; DB 1; Length 763;
Best Local Similarity 22.3%; Pred. No. 0.93;
Matches 105; Conservative 72; Mismatches 190; Indels 103; Gaps 20;
8 GPIDETPTPADLSAQGLEASANKSAB-----AORIAGAEAKPKE---SKTDSVERV 58

```

```

Db 336 GTLAESFRSMRDSLSLTDARATARAEDAREDAEQRAAEAREDAEARKDAQETA 395
QY 59 SILRSAV-----NALMSLADKLGIASSNSSSSSTSSA-----DVSSTTATAPT 101
Db 396 RALESAAADYEALTAVAD--GDLTRRVDSRDHDMARIGHALNDMLDDIETSVAAA-- 451
QY 102 PPPFTFDYKTOA-----QTAIDTFTSTSLADTQAALVLSQDAVTNIKOTAADEETAI 156
Db 452 ---TAFSDHVSDDAAQRVEADAGDAIDAGT---DVSTVADEISDGATE-----QTDRLHEV 500
QY 157 AAWEETKNADAVKYGQITELAKYASNDQATLDSLSGLKLTSPDLLQAALLQSVANNKAAE 216
Db 501 AGEVDDLSASAEVAETVASLADTAGQAASAVDD--GROATEDAVET---MDVDADAEAAA 557
QY 217 LKEMQDNVPVP--GKTPAIAQSLVDQTDATQIEKDGNAIRDAYFAGQNASGAVENA-- 273
Db 558 DAMDALDSEMDIGEIVDVIAIQINMLAL-----NASIEAARTGADGDFAVVADE 611
QY 274 -----KSNNSISNDSAKAIAIATAKQIABAQKKFPDPSILOEAEQWVIOAEKDLKNIK 327
Db 612 VKTLAEESRDAEDIESRLALQGVSDVADEMEATSDT--VSDGRATVGDATALDDV-- 668
QY 328 PADGSDVNPCTTGVGSKQGGSSIGSRVSMLLDDAENETASIIIMSGFROMIHMENTNP 387
Db 669 -----VSFVADITDTAAGEIRAA---TDRQHAASRVASAVDEV----- 703
QY 388 DSAQAQELAAQARA--AKAGD-----DSAAAALADAAQALEAALGK 428
Db 704 --AGISQETAQAQATAVADSAAATQDTLTSSVDDAAADLADRAAALLDLLAE 751
RESULT 10
HTR2 HALSA STANDARD; PRT; 764 AA.
AC P71410;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis
DE protein II) (MPP-II).
GN HTR2 OR HTRII.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FLX15;
RX MEDLINE=96323203; PubMed=8710852;
RA Zhang W., Brooun A., Mueller M.M., Alam M.;
RT "The primary structures of the Archaeon Halobacterium salinarum blue
RT light receptor sensory rhodopsin II and its transducer, a methyl-
RT accepting protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:8230-8235(1996).
RN [2]
RP FUNCTION.
RX PubMed=9515936;
RA Hou S., Brooun A., Yu H.S., Freitas T., Alam M.;
RT "Sensory rhodopsin II transducer HtrII is also responsible for serine
RT chemotaxis in the archaeon Halobacterium salinarum.";
RL J. Bacteriol. 180:1600-1602(1998).
CC -!- FUNCTION: Transduces signals from the phototaxis receptor sensory
CC rhodopsin II (SR-II) to the flagellar motor. Responds to light
CC changes through the variation of the level of methylation. Also
CC acts as a chemotransducer.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


188 KKAQYEQDLAANKAEVERSLMRKPRPIYAK---LAQNKDL--RAIQANSDSQAY 242
 255 -AIRDAY-----FAGQNASGAVENAKSNNSINSDSAKAIATATKQIAEAKKFPD 305
 243 AAAREAYDKWARVQAANAANKAYEEALANTAKN-DQIKAEIAIQORSKA----- 295
 306 SPILQEAQOMVIOAEKDKLNKIPADGSDVPNPGTTVGSKQOGSSIGSIRVSMILLDAEN 365
 296 -----DYEAUKAQYEQDKLAAQAQAGNAENADYQAKKAYEQELARV-----QAAN 340
 366 ETASILMSGFQMIHMENTEN-----PQSAQAQELAAQARAAXAAGDDSAALADAKA 421
 341 AAAK---QAYEQALAAANSKAKNAQITAEALQON--AQAKA-----DYEAUKAQYQKD 388
 422 LEAA-LGKAGQOQQGILNALG-----QTASAAVWSA 450
 389 LAAQAQSGNAANEADYQEKLAAYAEKELARVQAANAQA 427

RESULT 12

FLJB_SALTY STANDARD; PRT; 505 AA.
 ID FLJB_SALTY STANDARD; PRT; 505 AA.
 AC P52616; P97159;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phase-2 flagellin.
 GN FLJB OR H2 OR STM2771.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SL 375;
 RX MEDLINE=95325331; PubMed=7541401;
 RA Vanegas R.A., Joys T.M.;
 RT "Molecular analyses of the phase-2 antigen complex 1,2,. . of
 Salmonella spp.";
 RL J. Bacteriol. 177:3863-3864 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856 (2001).
 RN [3]
 RP SEQUENCE OF 1-37 FROM N.A.
 RX MEDLINE=62049491; PubMed=6271461;
 RA Silverman M., Zieg J., Mandel G., Simon M.;
 RT "Analysis of the functional components of the phase variation
 system.";
 RL Cold Spring Harb. Symp. Quant. Biol. 45:17-26 (1981).
 RN [4]
 RP SEQUENCE OF 482-505 FROM N.A.
 RC STRAIN=SU2353;
 RA Mingorance J., Tanaka S., Tomimaga A., Enomoto M.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Flagellin is the subunit protein which polymerizes to
 form the filaments of bacterial flagella.
 CC -!- MISCELLANEOUS: Individual salmonella serotypes usually alternate
 between the production of 2 antigenic forms of flagella, termed
 phase-1 and phase-2, each specified by separate structural genes.
 CC -!- SIMILARITY: Belongs to the bacterial flagellin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; U17177; AAC43354.1; --
 EMBL; AE008826; AAL21657.1; --
 EMBL; V01370; CAA24655.1; --
 EMBL; D26168; BAA05156.1; --
 DR StvGene; SG10564; fljB.
 DR InterPro; IPR001029; Flagellin C.
 DR InterPro; IPR001492; Flagellin N.
 DR Pfam; PF00700; Flagellin C; 1.
 DR Pfam; PF00669; Flagellin N; 1.
 DR PRINTS; PR00207; FLAGELLIN.
 DR ProDom; PD000316; Flagellin C; 2.
 KW Flagellum; Complete proteome.
 FT INIT_MET 0 BY SIMILARITY.
 FT CONFLICT 37 I -> S (IN REF. 3).
 SQ SEQUENCE 505 AA; 52404 MW; 3A5CC404AF7AF88B CRC64;

Query Match 6.7%; Score 147; DB 1; Length 505;
 Best Local Similarity 20.8%; Pred. No. 0.68; Indels 122; Gaps 19;
 Matches 92; Conservative 62; Mismatches 166;
 QY 23 AQLLEASAAKSAAPQRIAG-----AEAPKESKTSVSVERWSILRSVAVNALMSLADKL 75
 DB 32 SSGLRINSKDDAAGQAIAANRFTANIKGLTQASRNANDGISIAQTTEGALNEINNLRV 91
 QY 76 GIASSSSSTSRSAVDSTATAP-----TPPPTFDYKTAQATYDTFTS- 126
 DB 92 RELAVQSANSTNSQSDLDLSIQAEITQRLNEIDRVSGQTQFNGVKVLAQDNTLTQVGAND 151
 QY 127 --LADIAQALVSLQ---DAVT---NIKDTAAT-----DEBTAAAE 159
 DB 152 GEITIDILKQINSQTLGLDSLNVKQKVDVDTAVTTKAYANNGTLLDVGSLDAAIKAAAT 211
 QY 160 WETKNADAVKVGQITELAKYASDNQAILDSLGLKLTSTFD-----LLQ 201
 DB 212 GGTNGTASVTGGA-----VKFDADNNKYFVTIGFTGADAANKGYEVNVTADGTVTLLA 266
 QY 202 AALLQSVANNKNAELLKEMQDNFV---PGKTPAQSLVDQDTATATOI-----EKD 252
 DB 267 GAKTTPAGATTKTEVQELKDTFAVVSADAKNALTAGG-VDATDANGAELVKMSYTDKN 325
 QY 253 GNAR-----DAYFAG--QNASGAVENAKSN-----NSISNIDSAKAAI 289
 DB 326 GKTEGGYALKAGDKYVADYDEATGAIK-AKTSYTAADGTTKTAANQLGGVD-GKTEV 383
 QY 290 AT-----AKTQIAEAKKFPDPSPIQEAQOMVIOAE---KOLKNIKPA 329
 DB 384 VTIDGKTYNASKAGHDFKAQPELAEEAAKTENP-LQKIDAALAOVDAALRSLDGLAVQNR 442
 QY 330 DGSVDVENPGTTVGSKQOGSSI 351
 DB 443 FNSAITNLGNTVNNLSEARSRI 464

RESULT 13

HTR6_HALN1
 ID HTR6_HALN1 STANDARD; PRT; 778 AA.
 AC Q9HR92;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Halobacterial transducer protein Vi.
 GN HTR4 OR HTPVI OR VNG806G.
 OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]


```

RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Danon M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Feck R.F., Pohlischroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction (by similarity).
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005022; AAG19266.1; -.
DR F84237; F84237.
DR HSSP; F02942; IQU7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR004090; Me_chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDCUR.
DR SMART; SM00304; HAMP; 2.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS01111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 2.
DR Transducer; Transmembrane; Complete proteome; Repeat.
FT DOMAIN 1 26 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 27 47 POTENTIAL.
FT DOMAIN 48 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 317 POTENTIAL.
FT DOMAIN 318 778 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 318 370 HAMP 1.
FT DOMAIN 412 465 HAMP 2.
FT DOMAIN 484 720 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 778 AA; 82076 MW; D752278727AFA79 CRC64;

Query Match 6.6%; Score 146; DB 1; Length 778;
Best Local Similarity 23.5%; Pred. No. 1.3;
Matches 103; Conservative 63; Mismatches 166; Indels 106; Gaps 18;

QY 20 DLQAQGLEAARAKSAQAQRIAGAAKPKESKTD-----SVERWSILRSA----- 64
Db 378 DAFAARSEAAQADAAEAQAQ-ALFAARERSEQAQRRLTAAFASETMVAAGDLTVR 436
QY 65 VNALMSLADKLGTASSNSSSSSTSRSDVSDTATPTPPPTFDYKTAQATYDIFTS 124
Db 437 LDADVQAAMADIAAAYNEWAADMEATIDVVAFA-----DEVATASTDASD----- 483
QY 125 TSLADIQAALVSLQAVTNIKDTAA--TDETAIAAEWETKNADAVKVGAIQITELAKYAS 182
Db 484 -SAAAVQTGRDVSADVGRIDRAADQDQLEAVASETDEMSATIEEVAASAOVAB-TS 541
QY 183 DNQAILDSGLKLTDFLLQALLOSANNKAAELLKEMQNPVPGKTPALAQSLVDQT 242
Db 542 QRAAALGDG-----QAAQDAVAQ-----LEEIED-----ETQA-AATAVDL 579
QY 243 DATATQIEKGNATRD-----AYFAGQNASGAVENAKSNNSINISDAKA 287
Db 580 EAKMSEIETIVAAITDIAEQTNMLALNANIEAARADQDGGFAVVADEVKDLA--DESKA 637

```

```

QY 288 AATATAQTIAEAKQKFPDPILOAEQWVIOAEKDLKNIKPADGSDVNPFTVGGSKQQ 347
Db 638 SAAETIALVAEYR-----AQTETSVAAMDRIQSRVSDGVE-----TVSETERS 680
QY 348 GSSIGSIRVSMLLDAENETASILMSGFQMTMPTENPDQAAQCELAQAARAKAAG 407
Db 681 LSEIAG-RI-----AEADTG-----VQETSNAMDDQAAASVSDVTTAVGDVAALG 723
QY 408 DDSAAALADAQKALEAA 425
Db 724 EETATEAESTADAAAEQA 741

RESULT 14
ID PAC_STRMU STANDARD; PRT; 1565 AA.
AC P11657;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fac protein precursor.
GN PAC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]_TaxID=1309;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 39-48.
RC STRAIN=MT8148 / Serotype C;
RX MEDLINE=89343654; PubMed=2761390;
RA Okahashi N., Sasakawa C., Yoshikawa M., Hamada S., Koga T.;
RT "Molecular characterization of a surface protein antigen gene from
RT serotype c Streptococcus mutans, implicated in dental caries."
RL Mol. Microbiol. 3:673-678(1989).
CC -!- FUNCTION: SURFACE PROTEIN ANTIGEN IMPLICATED IN DENTAL CARIES.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -!- SIMILARITY: STRONG TO AGGLUTININ RECEPTOR OF S.SANGUIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14490; CAA32652.1; -.
DR PIR; S04729; S04729.
DR PDB; 1JWM; 17-JUL-02.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAme; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Antigen; Cell wall; Peptidoglycan-anchor; Repeat; Signal;
KW Dental caries; 3D-structure.
FT SIGNAL 1 38
FT CHAIN 39 1535 PAC PROTEIN.
FT PROPEP 1536 1565 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 219 464 3 X TANDEM REPEATS, ALA-RICH.
FT SITE 1532 1536 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1535 1535 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1565 AA; 170781 MW; 4C3B05C809D0C32A CRC64;

Query Match 6.6%; Score 145; DB 1; Length 1565;
Best Local Similarity 23.6%; Pred. No. 3.4;
Matches 107; Conservative 71; Mismatches 202; Indels 74; Gaps 22;

QY 11 DETERPPADLSAQGLE-ASAANKSAEAQRIAGAAKPKESKTDSEVERWSILRSVNALM 69
Db 39 DETTTTSDVDTKVGVGTGNPATNIPQAQGSASKAEQSQTKLE-----RQWHTTIE 90
QY 70 SLADKLGIASSNSSS-----STSRSDVSDTATAP-----TPPPPTFDYKTAQATYDITF 122

```

Db 91 VPKTDLQAAKDAKSGVNVQDADVNGKTVKTPBEAVQKETEIKEDYTKQAEIDIKT-- 148
 Qy 123 TSTSLADTQAAVLVSQDAVTNK-DTAATDEB-----TAIAAEWETKN-ADAVKVGQAQIT 175
 Db 149 TDQYKSDVAH-----EAEVAKTKAKNOATKEQYKDMAHKAVERINAANAASAKTAYEA 204
 Qy 176 ELAKVASDNOALDLSGLKLTSDLLQALLOSVAANNKAAELLKEMQDNPPVPGKTPAIA 235
 Db 205 KLAQY-----QADLAAVQKTNAN--QAYQKALAYQ--AEUKRVQEAN-----AAKAAY 252
 Qy 236 QSLVDQDTATAPQIEKDGNAIRDAYFAGQNASGAVEN-----AKSNNSISNIDSAKAA--- 288
 Db 253 DTAVAANNAKNTAIAANEIEIK-----RNATAKAEYETKLAQYQAEUKRVQEANAE 307
 Qy 289 -----IATAKQTIAEAKKFPSPILQAEQVWIOAEKDLKNK-PAGDSVFN-----P 337
 Db 308 DYQAKLTAYQTELARVQKANADAKATYEA-----AVAAANNAKVAALTAENTAIKQENENAK 363
 Qy 338 GTTVGSGSQQSSIGSIRVSMLLDDAENETASILMSGFROMHMENTENPDSQAQQAELA 397
 Db 364 ATYEAALQYEAADLAUVKKNANNEADYQAK-----ITAYQTELARVQKANADAKAYEAAV 420
 Qy 398 AQARAKAA--GDSAAAL-ADAQKALBAALGK 428
 Db 421 AANNAANNAALTAENTAIKKRNADAKADYEAKLAK 454

RESULT 15

SYEP_DROME STANDARD; PRT: 1714 AA.
 AC P28668; Q9VCF5;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bifunctional aminoacyl-tRNA synthetase [includes: Glutamyl-tRNA
 synthetase (EC 6.1.1.17) (Glutamyl-tRNA synthetase); Prolyl-tRNA
 synthetase (EC 6.1.1.15) (Proline-tRNA synthetase)].
 GN AATS-GLUPRO OR G53394.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92097547; PubMed=1756734;
 RA Cerini C., Kerjan P., Astier M., Gratecos D., Mirande M., Semeriva M.;
 RT "A component of the multisynthetase complex is a multifunctional
 RT aminoacyl-tRNA synthetase."
 RL EMBO J. 10:4267-4277(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97217441; PubMed=9063462;
 RA Cerini C., Semeriva M., Gratecos D.;
 RT "Evolution of the aminoacyl-tRNA synthetase family and the
 RT organization of the Drosophila glutamyl-prolyl-tRNA synthetase gene.
 RT Intron/exon structure of the gene, control of expression of the two
 RT mRNAs, selective advantage of the multi-enzyme complex."
 RL Eur. J. Biochem. 244:176-185(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Berkeley;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cheriy J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasok P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
 CC diphosphate + L-glutamyl-tRNA(Glu).
 CC -!- CATALYTIC ACTIVITY: ATP + L-proline + tRNA(Pro) = AMP +
 CC diphosphate + L-prolyl-tRNA(Pro).
 CC -!- SUBUNIT: Component of the multisynthetase complex which is
 CC comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
 CC monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
 CC arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
 CC proteins, p18, p48 and p43.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CLASS-I
 CC AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CLASS-II
 CC AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -!- SIMILARITY: Contains 6 WHEP-TRS domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M74104; AAA28594.1; -;
 CC EMBL; U59923; AAC47469.1; -;
 CC EMBL; AE003745; AAF56211.1; -;
 CC PIR; S18644; S18644.
 CC HSP; P00962; IGTR.
 CC FlyBase; FBgn005674; Aats-glupro.
 CC InterPro; IPR004526; GltX arch.
 CC InterPro; IPR000924; Glu tRNA-synt_1c.
 CC InterPro; IPR004046; GST-Cterm.
 CC InterPro; IPR004154; HGTF_anticon.
 CC InterPro; IPR004499; ProS_fam1.
 CC InterPro; IPR002314; tRNA-synt_2b.
 CC InterPro; IPR001412; tRNA-synt_1.
 CC InterPro; IPR002316; tRNA-synt_pro.
 CC InterPro; IPR006195; tRNA_ligase_II.
 CC InterPro; IPR000738; WHEP-TRS.
 CC Pfam; PF00043; GST_C; 1.
 CC Pfam; PF03129; HGTP_anticon; 1.
 CC Pfam; PF00749; tRNA-synt_1c; 1.
 CC Pfam; PF03950; tRNA-synt_1c_C; 1.

DR Pfam; PF00587; tRNA-synt_2b; 1.
DR Pfam; PF00458; WHEP-TRS; 6.
DR PRINTS; PR00987; TRNASYNTHGLU.
DR PRINTS; PR01046; TRNASYNTHPRO.
DR TIGRFAMS; TIGR00463; gltx_arch; 1.
DR TIGRFAMS; TIGR00408; proX_fam_I; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
DR PROSITE; PS00762; WHEP-TRS; 6.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Multifunctional enzyme; Repeat.
FT DOMAIN 170 754 GLUTAMYL-TRNA SYNTHETASE.
FT DOMAIN 755 800 WHEP-TRS 1.
FT DOMAIN 827 872 WHEP-TRS 2.
FT DOMAIN 901 946 WHEP-TRS 3.
FT DOMAIN 980 1025 WHEP-TRS 4.
FT DOMAIN 1055 1100 WHEP-TRS 5.
FT DOMAIN 1129 1173 WHEP-TRS 6.
FT DOMAIN 1174 1180 POLY-GLY.
FT DOMAIN 1207 1714 PROLYL-TRNA SYNTHETASE.
FT SITE 209 220 "HIGH" REGION.
FT SITE 438 442 "KMSKS" REGION.
FT BINDING 441 441 ATP (BY SIMILARITY).
FT CONFLICT 102 106 TSPLP -> DKSLA (IN REF. 3).
FT CONFLICT 233 234 VC -> AF (IN REF. 3).
FT CONFLICT 341 345 NTACA -> KYCVR (IN REF. 3).
FT CONFLICT 583 583 K -> R (IN REF. 3).
FT CONFLICT 692 692 L -> A (IN REF. 3).
FT CONFLICT 753 753 T -> S (IN REF. 3).
FT CONFLICT 802 802 T -> S (IN REF. 3).
FT CONFLICT 873 873 P -> T (IN REF. 3).
FT CONFLICT 887 887 G -> V (IN REF. 3).
FT CONFLICT 1201 1201 P -> PA (IN REF. 3).
FT CONFLICT 1461 1461 MISSING (IN REF. 3).
FT CONFLICT 1587 1587 G -> V (IN REF. 3).
SQ SEQUENCE 1714 AA; 189197 MW; 6PE8C58045E48A8C CRC64;

Query Match 6.6%; Score:145; DB 1; Length 1714;
Best Local Similarity 19.4%; Pred.No.3.7;
Matches 99; Conservative 75; Mismatches 193; Indels 142; Gaps 19;

QY 7 PGPI-----DETRPPADLSAQGLEASAAKSAEQAQRIAG-----A 43
DB 701 PSFIVLPSIPDGHKOVPTSGLVKNAPDAKATKASPSVSSGQASELDSQITQQGLVLR 760
QY 44 EAKPKSKTDSVERWSILRSVAVALMSLADKLGIASSNSSSTSRSDVDSTTAT-APT 102
DB 761 DLKSKRAAKQID-----VAVKKLLAL-----KADYKSATGKDWKPGQTSATSAPVP 807
QY 103 PPTFDDYKTAQATAYDTITFTSLADIQAALYSLODAVTNIKDTAATDEETAIA----- 157
DB 808 -----AASSSANDAV-----SVNASIVKQGLVRLDKGKASKPEIDAAVKTL 852
QY 158 ---AEMETKNADAVKVGQAITEAKYASDNQAI--LDSLGKLTSLFLLQALQLQSVANN 212
DB 853 ELKAAQVKLTGQDWKGTVPPTAPPSASAAPSVGNDSDVAQILS-----QITAQGD 903
QY 213 KAAELLKEMODNPVVFPGKTPAIAQSLVDQDTATQIEKDGNAIRDAYFAGQNASGAVEN 272
DB 904 KVRELKSAKADKATVDAAVKTLTSLKADYKAAAT-----GSDWKPGTTAPAPAAPVKV 956
QY 273 AKSNN-----SISNIDSKAIAIAKQIAEAQKFPDSPILQAEQWVQAEKOLKNI-- 326
DB 957 KQKNPDPAVLVTNLLNKIAQQGDKIQRLKSAKSEKSLVEAEVKLLALKTDYKSLTG 1016
QY 327 ---KPADGSDVNPNGTT-----VGGSKQOGSSIGSI----- 354
DB 1017 QEWKP--GTVAAPTWNVLDLTGG--DSGSDVGVSLVKIQAQGDKIRKLKSEKAAKNVI 1072
QY 355 --RVSMLL-----DDAENETASILMGFRQMIHFNTENPDSQAQCELA 397
DB 1073 DPEVKTLIALKGEYKYLISGKDWTPDAKSEPAVV-----KKEASPVSNASPAKDELTOEIN 1127

QY 398 AQARAAKAGDDSAASAAALADAAQKALEAAL 426
DB 1128 AQGEKVRAAKGNKAKEVIDAEVAKLLAL 1156

Search completed: March 24, 2004, 05:57:01
Job time : 10.3108 secs

GenCore version S.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:47:55 ; Search time 34.0475 Seconds
(without alignments)
4188.690 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204

Sequence: 1 MWNPIGPFDLTERTPPAD.....QGILNALGQIASAAVUSAGV 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertibrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2204	100.0	551	16	Q927H7
2	2201	99.9	715	2	Q46166
3	723	32.8	636	16	Q824X3
4	461.5	20.9	647	16	Q84627
5	197	8.9	973	16	Q8XD04
6	193.5	8.8	2310	16	Q8CMU7
7	192.5	8.7	2055	2	O85472
8	190	8.6	1327	2	Q9X7M2
9	190	8.6	2478	2	Q9L69
10	188	8.5	2478	2	Q9LCH2
11	186.5	8.5	790	16	Q8ZQF5
12	185.5	8.4	1822	2	Q07290
13	182	8.3	625	5	Q86RM2
14	181	8.2	2481	16	Q99QR6
15	180	8.2	621	5	Q86RM2
16	179	8.1	842	5	Q86RN1

17	176	8.0	625	5	Q86RN2
18	175	7.9	627	17	Q9HS86
19	175	7.9	855	5	Q86DH2
20	175	7.9	2345	16	Q8P6P6
21	173.5	7.9	2275	16	Q8NUL3
22	173.5	7.9	3360	16	Q88X86
23	173	7.8	2271	16	Q99QX4
24	172.5	7.8	812	16	Q8ZQ81
25	172.5	7.8	1128	5	Q26947
26	172	7.8	697	5	Q9NDJ0
27	172	7.8	856	5	Q86RM6
28	172	7.8	2016	5	Q9BIT0
29	170	7.7	993	16	Q8NUL3
30	169	7.7	1795	16	Q9LCH9
31	168.5	7.6	790	16	Q8ZNI3
32	168.5	7.6	1037	10	Q7XS51
33	168.5	7.6	2232	5	Q8IFX6
34	168	7.6	993	16	Q99QZ5
35	168	7.6	1545	16	Q9RDQ1
36	167	7.6	1033	10	Q7XVS0
37	166.5	7.6	803	5	Q86RN4
38	165.5	7.5	849	5	Q86R71
39	165.5	7.5	956	16	Q9LON7
40	165	7.5	732	5	Q86RM7
41	165	7.5	2283	2	Q8VQ99
42	164.5	7.5	639	5	Q86RM4
43	164.5	7.5	2167	2	Q845L8
44	164.5	7.5	9439	16	Q8CP76
45	164	7.4	6713	16	Q99U54

ALIGNMENTS

RESULT 1

Q927H7 PRELIMINARY; PRT; 651 AA.

AC Q927H7

DT 01-MAY-1999 (TREMREL. 10, Created)

DT 01-MAY-1999 (TREMREL. 10, Last sequence update)

DT 01-OCT-2003 (TREMREL. 25, Last annotation update)

DE CHLPN 76 kDa HOMOLOG 1 (CT622) (Hypothetical protein).

GN CPN0728 OR CP0728 OR CP0018 OR CP0756.

OS Chlamydia pneumoniae (Chlamydia pneumoniae).

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=83558;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CWL029;

EX MEDLINE=99206606; PubMed=10192388;

RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,

RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;

RT "Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.";

RL Nat. Genet. 21:385-389(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=AR39;

EX MEDLINE=20150255; PubMed=10684935;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,

RA White O., Hickey E.K., Peterson J., Utterback T., Berry K.,

RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,

RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,

RA Eisen J., Fraser C.M.;

RT "Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.";

RL Nucleic Acids Res. 28:1397-1406(2000).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=J138;

EX MEDLINE=20330349; PubMed=10871362;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,

RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;

RT "Comparison of whole genome sequences of *Chlamydia pneumoniae* J138

```

RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-183;
RA Geng M.M., Schumacher A., Muehlendorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RL other Chlamydia strains based on whole genome sequence analysis.";
DR Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001654; AAD18867.1; -
DR EMBL; AE002165; AAF37914.1; -
DR EMBL; AP002547; BAA98935.1; -
DR EMBL; AB017159; AAP98685.1; -
DR PIR; D72042; D72042.
DR PIR; E86581; E86581.
DR PHCI-2DPAGE; Q927H7; -.
DR TIGR; CP0018; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 651 AA; 68217 MW; 47AE6C3FF2FF0123 CRC64;

Query Match 100.0%; Score 2204; DB 16; Length 651;
Best Local Similarity 100.0%; Pred. No. 6.3e-98;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVNPIGPGIDETERTPPADLSAQGLEASAAANKSAEQAQRIAGAAKPKESKTDTSVERWSI 60
Db 1 MVNPIGPGIDETERTPPADLSAQGLEASAAANKSAEQAQRIAGAAKPKESKTDTSVERWSI 60
Qy 61 LRSVAVNALMSLADKLGIIASSNSSSSSTSRSDVDSTTATPTPPPTDDYKTAQTAYDT 120
Db 61 LRSVAVNALMSLADKLGIIASSNSSSSSTSRSDVDSTTATPTPPPTDDYKTAQTAYDT 120
Qy 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
Qy 181 ASDNQAILDSLGKLTSTFDLLQALLOSVAANNKKAELLKEMQNPVPGKTPAIQAQSLVD 240
Db 181 ASDNQAILDSLGKLTSTFDLLQALLOSVAANNKKAELLKEMQNPVPGKTPAIQAQSLVD 240
Qy 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Qy 301 KKFPDSPILQAEQMWIQAQEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLL 360
Db 301 KKFPDSPILQAEQMWIQAQEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLL 360
Qy 361 DDAENETASILMSGFQMIHNTENPDSSQAQELAAQARAARAAAGDSDSAAALADAQK 420
Db 361 DDAENETASILMSGFQMIHNTENPDSSQAQELAAQARAARAAAGDSDSAAALADAQK 420
Qy 421 ALEAALGKAGQGGQGIINALGQIASAAVVSAGV 452
Db 421 ALEAALGKAGQGGQGIINALGQIASAAVVSAGV 452

RESULT 2
Q46166 PRELIMINARY; PRT; 715 AA.
AC Q46166;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 76 kDa protein.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94156481; PubMed=7509320;
RA Perez-Melgosa M., Kuo C.-C., Campbell L.;

RT "Isolation and characterization of a gene encoding a Chlamydia
RL pneumoniae 76-kilodalton protein containing a species-specific
RN epitope.";
RP Infect. Immun. 62:880-886(1994).
RC EMBL; I23921; AAA23117.1; -.
DR PIR; I40729; I40729.
SQ SEQUENCE 715 AA; 76626 MW; 3F01C9A59DE964C5 CRC64;

Query Match 99.9%; Score 2201; DB 2; Length 715;
Best Local Similarity 99.8%; Pred. No. 9.8e-98;
Matches 451; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVNPIGPGIDETERTPPADLSAQGLEASAAANKSAEQAQRIAGAAKPKESKTDTSVERWSI 60
Db 257 LVNPIGPGIDETERTPPADLSAQGLEASAAANKSAEQAQRIAGAAKPKESKTDTSVERWSI 316
Qy 61 LRSVAVNALMSLADKLGIIASSNSSSSSTSRSDVDSTTATPTPPPTDDYKTAQTAYDT 120
Db 317 LRSVAVNALMSLADKLGIIASSNSSSSSTSRSDVDSTTATPTPPPTDDYKTAQTAYDT 376
Qy 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
Db 377 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 436
Qy 181 ASDNQAILDSLGKLTSTFDLLQALLOSVAANNKKAELLKEMQNPVPGKTPAIQAQSLVD 240
Db 437 ASDNQAILDSLGKLTSTFDLLQALLOSVAANNKKAELLKEMQNPVPGKTPAIQAQSLVD 496
Qy 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 497 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 556
Qy 301 KKFPDSPILQAEQMWIQAQEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLL 360
Db 557 KKFPDSPILQAEQMWIQAQEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLL 616
Qy 361 DDAENETASILMSGFQMIHNTENPDSSQAQELAAQARAARAAAGDSDSAAALADAQK 420
Db 617 DDAENETASILMSGFQMIHNTENPDSSQAQELAAQARAARAAAGDSDSAAALADAQK 676
Qy 421 ALEAALGKAGQGGQGIINALGQIASAAVVSAGV 452
Db 677 ALEAALGKAGQGGQGIINALGQIASAAVVSAGV 708

RESULT 3
Q824X3 PRELIMINARY; PRT; 636 AA.
AC Q824X3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN CCA00015.
OS Chlamydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J., Holtzaple E., Khouri H., Federova N.B., Carty H.A.,
RA Unayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
RA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavoil P.M.,
RA Fraser C.M.;
RT "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC):
RL examining the role of niche-specific genes in the evolution of the
RL Chlamydiaceae.";
DR Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016994; AAP04768.1; -.
DR TIGR; CCA00015; -.
KW Hypothetical protein; Complete proteome.

```



```

DR InterPro; IPR008969; Carboxypeptid_reg.
DR InterPro; IPR005003; Phage fiber.
DR InterPro; IPR005068; Phage fiber_2.
DR Pfam; PF03335; Phage fiber; 6.
DR Pfam; PF03406; Phage_fiber_2; 3.
KW Complete proteome.
SQ SEQUENCE 973 AA; 96316 MW; 776580D2A87E1B36 CRC64;

Query Match      8.9%; Score 197; DB 16; Length 973;
Best Local Similarity 22.9%; Pred. No. 0.13;
Matches 119; Conservative 63; Mismatches 208; Indels 130; Gaps 14;

QY 26 LEASAANKSAEQAORTAGAEAKPKSKTDSVERWSILRSVAVNALMSLADKLGIASSNSSSS 85
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 112 VEEVARNASAVQONTAAAKSASDASTSARBAATHATDAOSARAASATSAQOAS- 170
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 86 TSRSADVDTTATPTPPPTFDYKTOAQAYDTFTFTSLADIQALVSLQDAVINIK 145
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 171 ASSAGTASTKATEASKSAAAEKSSKAAATSAAGAKTSEI-----NNAVSCQSAATSAS 225
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 146 DTAATDETAIAEWETKNADAVKGAQITELAKYASDNQAILDSLGKLTGFDLLOALL 205
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 226 TAITKASEAASSA-----RDASAKSEAKSSETSAASSASSAATAAGNS-----AKAA 276
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 206 QSVANNKAAELLKEMQDNVPVPGKTPA-----IAQSLVDCTDATATQIEKD----- 252
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 277 KTSNTAKSSETAEQSAASAAGSKTAAALSASAASTASAGQASATAAGKSAESAASSA 336
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 253 -----GNATRDYFAGQNASGA---VENAKSNNSIS----- 280
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 337 STATTKAGEATEQSAASAASAKTSETNAKASSETSAESKTAASASSAASSASSAS 396
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 281 -----NIDSAKAAATATQIAE-----AQKFPSPILQEAQWVIAEKDLK 324
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 397 ASKDEATKQSAKSSATFTKATEAGSATAAQAQSKTAEASATRAETAAKRAEDIAS 456
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 325 NIKPADGSDVPNGTIVGGSKQGGSSIGSRVSMLLDDAENETASILMSG----- 374
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 457 AVALEDAS-----TTKGIQVLSATNSTSESIAATPKAVKAYELANGKYTAQDATT 510
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 375 FROMHMTNPDSQAAQOELAQRAAKAAGD-----D 409
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 511 QKGIQVLSNATNS*SEM-----LAATPKSVKAAAYDLANGKYTAQDATTAKGIQVLS 566
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 410 SAAALADAQKALEAALGKAGQOQGLNALQILASAAVVS 449
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 567 SASETLAATPKAVRAAND-----NANGRVSPSARKVN 597
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
Q8CMU7
ID Q8CMU7 PRELIMINARY; PRT; 2310 AA.
AC Q8CMU7; 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Streptococcus hemagglutinin protein.
GN SE2249.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016751; AAC05891.1;
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR008009; He_PIG.
DR Pfam; PF05345; He_PIG; 2.
KW Complete proteome.
SQ SEQUENCE 2310 AA; 232314 MW; AC96B49EAB754BB8 CRC64;

Query Match      8.7%; Score 192.5; DB 2; Length 2055;
Best Local Similarity 23.2%; Pred. No. 0.55;

QY 10 IDETERPPAD-LSAAGLE-ASAANKSAEQAORTAGAE-AKPKSKTDSVERWSILRSVAVN 66
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 964 LSSASATSDSASTQASEAS-SKLSSESTSTSTSDSASTSTSESTSESTSESTSEST 1023
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 67 ALMSLADKLGIASSNSSSSRSADVDTTATPTPPPTFDYKTOAQAYDTFTFTS 126
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1024 A--SLSESTSTSDSASTSESDNS-TSTSLSESTSTSLSESTSTSLSESTSTSL 1079
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 127 LAIQAAALVSLQDAV-TNIKD---TAATDETAIAEWETKNADAVKGAQITELAKYA- 181
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1080 VSDNSASTSLREBETSTSLSDSTSTSTSDSASTSTSESDSDSASTSLSGSSSTSV 1139
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 -----SDNQAILDSLGLTSPDLQ-----AALLOSVAANNKAAELLK 219
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1140 ASTSESASTSTSDNSNSTSTSLSESTSTSESTSTSTSDSASTSTSVSDNSASISLS 1199
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 220 EMQDNVPVPGKTPAIAQSL-----VDQTDATATQIEKDGNNAIRDYFAGQNASGA 273
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1200 ESTSTSVSDSTSTSEASSTSTSESDNSASTSLSESTSTSVSDS-----TSTSTSD 1254
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 274 KSNNSINIDSAKAAI-----ATAKQIAEAAQKFPDPSILQAEQ 314
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1255 STSTSVSDNSASTSLSESTSTSDSTSTSTSDSASTSTSVSES-----SSTSTSI 1308
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 315 MVIQAEKDLNIXIPADGSDVPNGTIVGGSKQGGSSIGSRVSMLLDDA----- 363
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1309 SLSTSDSDSKMSSTSSA---STSTSVSDSEASTSLSESTSTSLSDSTSTSDTSAS 1365
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 364 -----ENETASILMSGFROMHMTNPDSQAAQOELAAQRAAKAAGDSDSAAALADA 418
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1366 TSESDNSTSTSMESLSLSTSVSDSTSTSDSASTSTSVSDNSASTSLSDSTSTSDS 1425
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 419 QKA 421
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1426 TSA 1428
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
Q85472
ID Q85472 PRELIMINARY; PRT; 2055 AA.
AC Q85472; 1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Extracellular matrix binding protein (Fragment).
GN EMB.
OS Abiotrophia defectiva.
OC Bacteria; Firmicutes; Lactobacillales; Aerococcaceae; Abiotrophia.
OX NCBI_TaxID=46125;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VVS-47;
RX MEDLINE=99081722; PubMed=9864195;
RA Manganello R., van de Rijn I.;
RA "Characterization of emb, a gene encoding the major adhesin of
RT Streptococcus defectivus."
RL Infect. Immun. 67:50-56(1999).
DR EMBL; AF067776; AAD03320.1;
DR PIR; T31110; T31110.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR Pfam; PF04650; Ysirk signal; 1.
DR TIGRfams; TIGR01168; Ysirk_signal; 1.
FT NON_TER 2055 2055
SQ SEQUENCE 2055 AA; 215642 MW; 9699C11DDE93E2FD CRC64;

Query Match      8.7%; Score 192.5; DB 2; Length 2055;
Best Local Similarity 23.2%; Pred. No. 0.55;

```


RT "tn51-mediated insertional inactivation of the *fmeB* gene encoding a
RT cell wall-associated protein abolishes methicillin resistance in
RT *Staphylococcus aureus*.";
RL J. Antimicrob. Chemother. 45:421-431 (2000).

Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2.";
RL Nature 413:852-856(2001).
DR EMEL; AE008737; AAL19862.1; -;
DR InterPro; IPR008969; CarboxypepD reg.
DR InterPro; IPR005068; Phage fiber 2.
DR Pfam; PF03406; Phage_fiber_2; 1.
DR Complete proteome.
KW Complete proteome.
SO SEQUENCE 790 AA; 79050 MW; CB8E62739B80C9C3 CRC64;

Query Match 8.5%; Score 186.5; DB 16; Length 790;
Best Local Similarity 22.9%; Pred.No. 0.33;
Matches 113; Conservative 68; Mismatches 210; Indels 103; Gaps 18;

QY 8 GPIDETERTPPADLSAQGLEASAAANKSAEACRITAGAEAKPKESKTDVSVERNSILRSVNA 67
DB 91 GAMTEDDVREALURFFAMVEEVARQASEARNATAAGASBQASQTSAGO-----AAES 144
QY 68 LMSLADKLGIASSNSSSTSRSDVDSTTATAPTPPPPTDDYKTAQTAYDTIFTSTSL 127
DB 145 ATAAASAAGAADASATQAASASAESSAGTATK-----AGEASASAADTARTA 196
QY 128 ADTQAALVSIQDAVTNKTQATDDEETAIAAEWETKNADAVKVGQAQITELAKYASDNQAI 187
DB 197 AAASEAAAKTSEANADASRTAAGDSAAAAASAATAAQTSAERAGA--SETAAKTSETQA- 253
QY 188 LDSISGLKLTFDLLCQALLQSVANNKKAELLKEMQDNFVVPVPGKTPA----IAQSLVDQTD 243
DB 254 ASSAGDAGAGATAAASEKAAAS--AAE-AKTSETNAATSASTAAASATAAASSSASEAS 310
QY 244 ATATQIEKQGN-AIRDAYFAGQASGAVENAKNNSISNI----- 282
DB 311 THAASDTSASLAQSPAAGESATRAEEAAKRAEDIADVISEDASLTKKGIVQLSSAT 370
QY 283 DSAKAIAT-----AKTQIAPAAQKFF-DSPII-----QEA 312
DB 371 DSDSEALASTPKAVKTVMGVEVQTKAPLDSPALTGTPAPTPTTAAGIETATAAFVAAKV 430
QY 313 EQMVIQAEK---DLKNIKPADGSDVDNPGTGVGSKQCGSSIGSIRVSMLLDDAENETAS 369
DB 431 AQVUGSAPEALDITKELADALGND-PNFATTV-----LNKLAGKQPLDITLTALSG 480
QY 370 ILMGSRFQMIHMENTENPDSQAQELAAQARAAKAAGDSDSAAALADAAQKALEAALGKA 429
DB 481 KSVVDGLIYIGLNTID-----KAAGALPAGGTAVANRILVSRGALPALTGT 528
QY 430 -GQQQGIILNALGOI 442
DB 529 RGSDSGLI--MGEV 540

RESULT 12
Q07290
ID Q07290 PRELIMINARY; PRT; 1822 AA.
AC Q07290;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Epf* protein.
DE GEPF*.
OS Streptococcus suis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1307;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=1890;
RX MEDLINE=93328288; PubMed=8335363;
RA Smith H.E., Reek F.H., Vecht U., Gielkens A.L.J., Smits M.A.;
RT "Repeats in an extracellular protein of weakly pathogenic strains of

```

RT Streptococcus suis type 2 are absent in pathogenic strains.";
RL Infect. Immun. 61:3318-3326(1993).
DR EMBL; X71880; CAA50714.1; -.
DR PIR; S33441; S33441.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR Pfam; PF04650; YSIRK_signal; 1
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
SQ SEQUENCE 1822 AA; 192632 NW; 3838960C77641D7D CRC64;

Query Match      8.4%; Score 185.5; DB 2; Length 1822;
Best Local Similarity 22.8%; Pred. No.1;
Matches 119; Conservative

Qy 10 IDETERTPPADLSAQGLEASAAKSAFAQRIAGAAKPKESKTDV---ERWSILSAVN 66
Db 1072 IDNPNUTPDB-----KESAKVAEEAAKVATA-AIDKASTPDAVQVEDKGV--AAIN 1122
Qy 67 ALMSLADKLGIASSNSSSTSRGAD-----VDSTTAT-----APTPPPPTFDYKT 112
Db 1123 LIATKADAGVIAAKLADEIKKLEDKQAEAKAIDASTWTNEEKATAKKALQDVVKGKA 1182
Qy 113 Q-----AQTAVDTTFTSTSLADIQAALV-----SLQDAVTNIKOTAADEETAIAAE-WETK 163
Db 1183 ELEDAARVATNEIHEATTTTEKAKAAELAGKSLTDTGKEARDAVELAKOKELAKEAIRTE 1242
Qy 164 NADAVKGAQITELAKYA-----SDNOA-----ILDSLGKLTSPDL 199
Db 1243 EBEATKIVEKLAEDTRKATIEDNPNLSDKQAEIKKLTDAVAKTLATIRNADRTQ--- 1299
Qy 200 LQALILQSVANNKKAELLKEMQDNVPVFGKTPAIAQSLVDQDTATA-TQIEKDGNAIRD 258
Db 1300 -EAKAQALADLEKAKETQK-IADKAAIDRLTILVKGDELEATKQDAKNKIADKAAAAAKE 1357
Qy 259 AYTAGQWASGA-----VENAKSNNSIS-----NIDSAKAAIA----- 290
Db 1358 AIASNPNLNTDAEKKTFPTDAVDAEVAKANDAIASAATSADVQKEEDAGVAAIADVLDAK 1417
Qy 291 -TAKTQIAE----AOKKFPDSPILQEAQWVIOAEKDLKNIKPADGSDVPNPGTTVGGSK 345
Db 1418 QDAKNKIADKAAAAKEAIGSNPNLNTDAEKKTFPTDAVDAEVAKANDAIASAATSADV--QK 1475
Qy 346 QCGSSSTGSTRVSMI-----LDDAENETASILMSGPFQMHMFNTENPNPSQAQOELAAQARA 402
Db 1476 EEDAGVAAIAEDVLDAAKQDAKNKIAKESDAAKSAIDANPNLNTDAEKESAKKAVDADAKA 1535
Qy 403 AKAAAGDSDSAAALADAKALEAALGKAGQOQGIILNALGOIA 443
Db 1536 ATDAID--ASTSPVEAQSDKGVGSIAD--VLDAAKQDA 1572

RESULT 13
Q86RN6 PRELIMINARY; PRT; 625 AA.
AC Q86RN6;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DE 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Merozoite surface protein 3 alpha (Fragment).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
SEQUENCE FROM N.A.
RP SPRAIN=IEC-Brazil 67;
RC MEDLINE=22346636; PubMed=12458823;
RA Rayner J. C., Corredo V., Feldman D., Ingravallo P., Iderabullah F.,
RA Galinski M.R., Barnwell J.W.
RT "Extensive polymorphism in the plasmodium vivax merozoite surface coat
RL protein MSP-3alpha is limited to specific domains.";
RL Parasitology 125:393-405(2002).
DR EMBL; AF491946; AAC20877.1; -.
FT NON TER 1

```

```
FT NON TER 625 625
SQ SEQUENCE 625 AA; 65303 MW; ADA71E17C2914047 CRC64;

Query Match
Best Local Similarity 8.3%; Score 182; DB 5; Length 625;
Matches 102; Conservative 66; Mismatches 223; Indels 82; Gaps 12;

Qy 23 AQGLEAAGAKSAFAQRIA-GAEAKPKESKTDVSVE-----RWSILRS 63
Db 81 AEEAABAKQAATAEKAAADAAEAGKENDLDDVKSQVKIAEEASTNAKKNKTEAIVE 140

Qy 64 AVNALMSL-----ADKLGIASNSSSTSSADVDSTATAPPPPTFFDDYKTOA 114
Db 141 IVKAVVAKEEAQKASDEAKACEKAQKAHAKQASDTTKTVERFKTNAEAAAKNAKEKA 200

Qy 115 QTAVDYTFSTSLADIAQALVSLODAVNIKDTAATBETAIAAEWETKNADAVKVAQI 174
Db 201 ENANKAATEASANELSVAKQKATDAETAAKDAKKEQVKAELAE-----VAKAKVAKEE 255

Qy 175 TELA-KYASDNOAILDSIGKILTSFDLLQALLOSANNKAAELLKEMQDNVVPVGGTPA 233
Db 256 ADAQKKAEEAKKIVDKIADKSEVPEAQKAAEFATETVKATTAATEAGKNAQEAESPE 315

Qy 234 IAQSLVDOTDATAQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKALATAK 293
Db 316 EAERKAETSADVKGAD-----AAEKAAGEAKKASIETEI-----AVEVAK 356

Qy 294 TOIAEAQKFFPDSPILOEAEQWVIOAEKDLKNIKFPA-----DGSVNPNGTTVGGSK 345
Db 357 AEVLNAEVK----KTAQAEKDAETAEKQAEKAKAAEAEKTHEKAKGVESTKAHSDK 412

Qy 346 QCGSSIGSIRVSMLLDDAENETASILMSGFRQIMHFMNTENPDSQAAQELAAQARAKA 405
Db 413 AQQENKNAKAS---EEAENRAVDALBEAYAVEAHLARTKNAEBSAKSATSMDSELEKAKE 469

Qy 406 AGDSSA-----AALADQAQKALEAALGKAGQOQILNALGQIASAA 446
Db 470 BAIDAANIAHQWIKATQANTIAKEK--EAAKVAEKAQTAANVVKDKAAGA 520

RESULT 14
Q99QR6 PRELIMINARY; PRT; 2481 AA.
AC Q99QR6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE FmtB protein
GN FmtB (MRP) OR FmtB OR SAV2160 OR SA1964.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPCCIRS=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyana A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB56322.1; -.
DR EMBL; AP0033136; BAB43253.1; -.
DR PIR; D90011; D90011.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.

DR GO:0016774; F:phosphotransferase activity, carboxyl group. .; IEA.
DR GO:0008152; P:metabolism; IEA.
DR GO:0016310; P:phosphorylation; IEA.
DR InterPro; IPR000890; Acetate kin.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
DR PROSITE; PS01075; ACETATE_KINASE_1; 1.
KW Complete proteome.
SQ SEQUENCE 2481 AA; 263765 MW; EIEAAB99B81665E8 CRC64;

Query Match
Best Local Similarity 8.2%; Score 181; DB 16; Length 2481;
Matches 127; Conservative 91; Mismatches 204; Indels 126; Gaps 25;

Qy 4 PIGPGPIDE-----TERTPEADLSAQGLEASAANKSA-EAORTAGAEAKPKESKTDS--- 54
Db 763 PLNPDTTNEEVAAIERINAAKVS--GVKAIEATTTAODLERVKNBEIFKIENITDSTQT 820

Qy 55 -----VERWSILRSVAVNALMSLADKLGIASNSSSTSSRSAD-----VDSSTATAPT 101
Db 821 KMDAYKEVRGAATARKAQAATVSNATDEEVAENAAVDAQAQTEGLHDIOVVKQOEVDAT 880

Qy 102 PP-----PPTFFDYKTOAQPAYDT-----IFTSTSLADIAQALVSL--- 137
Db 881 KAKVLDKINAIQOAKVKPAAD--TEVENAYNTRKQEIQNSNASSTEEKEAAVTELDK 937

Qy 138 -QDAVINIKDTAATD-----EETAIAAEWETKNADAVKGAQITELAKVASDNOAILDS 190
Db 938 KQEARITNL-DAANTNSDVTTAKONGIAAINQVQAATPKGSDAK-AETIAQKASERKTAIEA 995

Qy 191 LGKLT-----SPDLQALLOSANNKKA-----AELLKEMQDNVVPVPGKT--PAI 234
Db 996 MNDSTTEEQAAKDKVDQAVVTANADIDNATANTVDNKTTEATIAATTPDANKVPA 1055

Qy 235 AQSLVDOTDATAQIE-KQGNAIRDAYFAGQN-----ASGAVENAKSNNSI----- 279
Db 1056 KQAIADKQVQAEATAIDANNSTTEKEAAKQVQTEKTAADAADAIDAAHSNVEVEAAKNAE 1115

Qy 280 -----SNIDSKAIAIAT-----AKTQIAEAKKKFPDPSPILOEA--EQWVIOAEKD 322
Db 1116 IAKIEATQPATTTKONAKQAIATKANERKTAIACTODITAEETIAANADVNDVNTQAN-- 1173

Qy 323 LKNIKPADG-SDVNPFGTTVGGSKQCGSSIGSIRVSMLLDDAENETASILMSGFRQIMHM 381
Db 1174 -SNIEAANSQNDVDQAKTGTSETSIDQVTPVNNKAT-----ARNEITAILNNKIQEIQAT 1227

Qy 382 FNTENPDSQAAQOEELAAQARAQAAGDSDSAAA-----ALADAQKALEAALGKAGQOQG 434
Db 1228 PDATDEEKQADAE--ANTENKANQAISAATTNAQVDEAKANAEEAIAINAVTPRVVKKQA 1285

Qy 435 ILNALGOI 442
Db 1286 AKDEIDQL 1293

RESULT 15
Q86RM2 PRELIMINARY; PRT; 621 AA.
AC Q86RM2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Merozoite surface protein 3 alpha (Fragment).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sri Lanka I;
RX MEDLINE=22346636; PubMed=12458823;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:53:56 ; Search time 12.3321 Seconds
(without alignments)
3525.629 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204
Sequence: 1 MWNPIGPGIDETERTPPAD.....QGILNALQIASAASVAGV 452

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.78:*
1: Pirl.*
2: Pirl.*
3: Pirl.*
4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2204	100.0	651	2	D72042 conserved hypothetical
2	2204	100.0	651	2	D86581 CHLPN 76 kDa homol
3	2201	99.9	715	2	I40729 hypothetical 76K p
4	461.5	20.9	647	2	G71490 hypothetical prote
5	197	8.9	971	2	B30835 probable tail fibe
6	197	8.9	973	2	C85693 probable membrane
7	192.5	8.7	2055	2	T31110 extracellular matr
8	185.5	8.4	1822	2	S33441 EF protein - strep
9	184	8.3	1122	2	G64887 probable tail fibe
10	181	8.2	2481	2	D30011 FmtB protein [mpo
11	176.5	8.0	545	2	E94327 Htr7 transducer [i
12	176.5	8.0	545	2	T46811 halobacterial tran
13	175	7.9	627	2	F84194 Htr14 transducer [
14	173	7.8	2271	2	F30073 hypothetical prote
15	172.5	7.8	1128	2	T30296 R27-2 protein - Tr
16	168.5	7.6	2232	2	T34434 hypothetical prote
17	168	7.6	993	2	C30072 hypothetical prote
18	164	7.4	5713	2	B89921 hypothetical prote
19	163.5	7.4	2155	2	A2742 conserved hypothet
20	163.5	7.4	2155	2	C97523 hypothetical prote
21	160.5	7.3	544	2	T44938 transducer protein
22	160	7.3	536	1	A47190 transducer protei
23	160	7.3	536	2	E84318 Htr1 transducer [i
24	158.5	7.2	4776	2	E95206 cell wall surface
25	157	7.1	641	2	C82206 methyl-accepting c
26	156	7.1	810	2	T46810 halobacterial tran
27	156	7.1	810	2	F84327 Htr5 transducer [i
28	156	7.1	2186	2	H89960 hypothetical prote
29	154.5	7.0	1238	2	T03465 probable exonuclea

ALIGNMENTS

RESULT 1

D72042
Conserved hypothetical protein CP0018 [imported] - Chlamydia pneumoniae (strains CWI
N; Alternate names: chlpn 76 kDa homolog 1 (ct622); hypothetical protein CPN0728
C; Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C; Date: 23-Apr-1999 #sequence_revision-23-Apr-1999 #text_change 11-May-2000
C; Accession: D72042, D87623
R; Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A; Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.
A; Reference number: A72000; MUID:99206606; PMID:10929388
A; Accession: D72042
A; Molecule type: DNA
A; Residues: 1-651 <ARN>
A; Cross-references: GB:AE001654; GB:AE001363; NID:94377031; PIDN:AAD18667.1; PID:9437703
A; Experimental source: strain CWL029
R; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A; Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A; Reference number: A81500; MUID:20150255; PMID:10684935
A; Accession: D81623
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-651 <REA>
A; Cross-references: GB:AE002165; GB:AE002161; NID:97188948; PIDN:AAF37914.1; PID:9718894
A; Experimental source: strain AR39, HL cells
C; Comment: This sequence was originally identified as homologous to part of a sequence
PIR:H71490.
C; Genetics:
A; Gene: CPN0728; CP0018

microtubule-associ
hypothetical prote
tail fiber protein
phage lambda-relat
lmp1 protein - Myc
large repetitive p
related to transcr
nucleolar phosphop
probable autotrans
intermediate filam
transducer protein
probable secreted
surface antigen sp
Htr2 transducer [i
transducer protein
hypothetical prote

Query Match 100.0%; Score 2204; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 7.2e-110;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MWNPIGPGIDETERTPPADLSAQGLEASAAKSAQAQIAAGAEKPKESKTSVERWSI 60
Db 1 MWNPIGPGIDETERTPPADLSAQGLEASAAKSAQAQIAAGAEKPKESKTSVERWSI 60
Qy 61 LRSVNALMSLADKLGIATSSNSSSTSRSDVSTTATATPPPTDDYKTOATAYDT 120
Db 61 LRSVNALMSLADKLGIATSSNSSSTSRSDVSTTATATPPPTDDYKTOATAYDT 120
Qy 121 IFSTSLADLQALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKGAQITELAKY 180
Db 121 IFSTSLADLQALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKGAQITELAKY 180
Qy 181 ASDNQAILDSLGKITSFDLLQALLOQSVANNKAAELLKEMQDNFVVPKTPAQAQSLVD 240
Db 181 ASDNQAILDSLGKITSFDLLQALLOQSVANNKAAELLKEMQDNFVVPKTPAQAQSLVD 240

C;Genetics:
A;Gene: C7622

Query Match 20.9%; Score 461.5; DB 2; Length 647;
Best Local Similarity 29.7%; Pred. No. 2.6e-17;
Matches 138; Conservative 86; Mismatches 192; Indels 49; Gaps 13;

QY 2 VNPFGPDPIDETERTPPADLSAQLSAGLEASAAKAGAEQRIAGABAKPKESKTSVERMSIL 61
DB 15 MNPILINGI-----ASNSETKESTKSEA-----SPSASSSVSSWSPL 52

QY 62 RSVAALMSLADKLGIIASSNSSSTSRs-ADVDSSTATATATPPPPFDYKTAQATYDT 120
DB 53 SSAAKHALISLRD--AILNKNSSPTDSLSQLEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110

QY 121 IPTSTSLADIQAALVSLQDAVTNIKDTAATDEE-TAJAABWETKNADAVKVGQAITELEAK 179
DB 111 LENATTLAEYETKADLMAALQDMERLAKOKAEVTRIKEALQEKQ-----EVIDKLNQLVK 166

QY 180 YASDNQAILDSLGKLTSTFDLLQALLOSVAANNKAAELLKEMQDNVPVPGKTPALAQSLV 239
DB 167 LEQONQTLKETITTTDSADQIPALINSQLEINKNSADQIIKDLEGONI---SYEAVLITNAG 223

QY 240 DQDTATATQIEKQGNARTRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATKTQIAEA 299
DB 224 EVIKASSEAGIKLQALQSLVDAGDSQAAVLQACQNNSPDNIAATKKLIDAAETKYNEL 283

QY 300 QKK---FPDPSPILOEAEQMVIAEKDLKNTKPADGSDVPNPGTTVGGSKQGGSSIGSI-- 354
DB 284 KOEHTGLTDSPLVKABEQISOAKDIOETIKPS-GSDIPIVGPS-GSAASAGSVAAGLKS 341

QY 355 -----RVSMILLDDAENETASTILMSGFRQMIMHNTENPDQAAQOEELAAQARAAK---AA 406
DB 342 SNNSGRISLILLDDVNEMAALMAGGERSMIBQFNVPNPATAKELQAMEAQITAMSDQLVG 401

QY 407 GDSAAALADAQKALFAALGKAGQOQGIINALGOIASAAVVSAG 451
DB 402 ADGELPAEIQAKDALAQAL-KPSTDGLATAMGQVAFAPAAKVG 445

RESULT 5

probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: B90835
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-971 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA035073.1; PID:g13361114; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC51650

Query Match 8.9%; Score 197; DB 2; Length 971;
Best Local Similarity 22.9%; Pred. No. 0.0049;
Matches 119; Conservative 63; Mismatches 208; Indels 130; Gaps 14;

QY 26 LEASAANKSAEQAQRIAGAEAKPKESKTSVERWSILRSVAVNMLSLADKLGIIASSNSSSS 85
DB 110 VEEVARNASAVAQNTAAAKKSADASTSAREAAHTATDAADSARAASATAGNS-----AKAA 168

QY 86 TSSADVSTTATATPTPPPTDDYKTAQATYDTFTSTSLADIQAALVSLQDAVTNIK 145
DB 169 ASSAGTASTKATEASKAAAEESKSAATISAGAATSET-----NAAVQQAASATAS 223

QY 146 DTAATDEETAIAAEWETKNADAVKVGQAITELEKAYASDNQAILDSLGKLTSTFDLLQALL 205

DB 224 TATTKASEAASA-----RDASAKAEAKSSETSAASSASAASSATAAGNS-----AKAA 274

QY 206 QSVANNKKAELLKEMQDNVPVPGKTPA-----IAQSLVDQTDATATQIEKD----- 252

DB 275 KTSETNAKSETAEQASASAAAGSKTAAALSASAASTAGQASASATAAGKSAESAASA 334

QY 253 -----GNAIRDYFAGQNASGA-----VENAKSNNSIS----- 280

DB 335 STATTKAGEATEQASAAASASAAKTSETNAKASETSABSSKTAASASSASAASSASSAS 394

QY 281 -----NIDSAKAAIATKTQIAE-----AQKFPDPSILOEAEQMVIAEKDLK 324

DB 395 ASKDEATROQASAAKSATTAATKATKAAGSATAAQAOSKSTAESAAATRAETAAKRAEDIAS 454

QY 325 NIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMILLDDAENETASTILMSG----- 374

DB 455 AVALEDAS-----TTKKGIVQLSSATNSTSESIAATPKAVKAAVELANGKYTAQDATTA 508

QY 375 FROMIMHNTENPDQAAQOEELAAQARAAGAAGD-----D 409

DB 509 QKGIVQLSNATNSTSEM---LAATPKVKAAYDLANGKYTAQDATTAQKGIVOLSSATN 564

QY 410 SRAAALADAKALEALGKAGQOQGIINALGOIASAAVVS 449

DB 565 SASSETLAATPKAVKAAND-----NANGRVPSARKVN 595

RESULT 6

probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli (strai
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: C85693
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85693
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-973 <STO>
A;Cross-references: GB:AE005174; NID:g12514847; PIDN:AGS6007.1; GSPDB:GN00145; UWGP:Z19
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1918

Query Match 8.9%; Score 197; DB 2; Length 973;
Best Local Similarity 22.9%; Pred. No. 0.0049;
Matches 119; Conservative 63; Mismatches 208; Indels 130; Gaps 14;

QY 26 LEASAANKSAEQAQRIAGAEAKPKESKTSVERWSILRSVAVNMLSLADKLGIIASSNSSSS 85
DB 112 VEEVARNASAVAQNTAAAKKSADASTSAREAAHTATDAADSARAASATAGNS-----AKAA 170

QY 86 TSSADVSTTATATPTPPPTDDYKTAQATYDTFTSTSLADIQAALVSLQDAVTNIK 145
DB 171 ASSAGTASTKATEASKAAAEESKSAATISAGAATSET-----NAAVQQAASATAS 225

QY 146 DTAATDEETAIAAEWETKNADAVKVGQAITELEKAYASDNQAILDSLGKLTSTFDLLQALL 205

DB 226 TATTKASEAASA-----RDASAKAEAKSSETSAASSASAASSATAAGNS-----AKAA 276

QY 206 QSVANNKKAELLKEMQDNVPVPGKTPA-----IAQSLVDQTDATATQIEKD----- 252

DB 277 KTSETNAKSETAEQASASAAAGSKTAAALSASAASTAGQASASATAAGKSAESAASA 336

QY 253 -----GNAIRDYFAGQNASGA-----VENAKSNNSIS----- 280

DB 337 STATTKAGEATEQASAAASASAAKTSETNAKASETSABSSKTAASASSASAASSASSAS 396

QY 281 -----NIDSAKAAIATKTQIAE-----AQKFPDPSILOEAEQMVIAEKDLK 324

Db	118	EEERA-----EAERAREKEAQKAEER---QTAEASAKQDARERSAEIEQLAADLESQ	169
Qy	72	ADKLGIASSNSSSTGRSADVDSITTATPTPPPT-FDDYKTAQATAYDTI-----FTST	125

Db 170 ATEVG-ATLEAASDGLTARVDATTDAENAEIAEVAATVNDMLTMTERTIDEIQGFSTNVT 228
QY 126 SLADIQAALVSLQDVTNIKDT---AATDEE---ATAAEWETKNADAVKGAQITE 176
Db 229 ASRETAGAKETODASQTVESVQIEAGTDQREQLSVAEEMDSYATVEEVAATAQS 288
QY 177 LAKYASDNQALDLSLGLT-----SFDLLQAALQSVANNKAAELKEMQDNVPVPGK 230
Db 289 VADTAADTTDVA-TAGQTAEDDAIDAIDAVQETMTTANVDALEDLTTEIDD----- 340
QY 231 TPAIAQSLVDQDTATATQIEKQGNNAIRDYFAGQNASG-----AVENAKNSNSI 279
Db 341 IAEILSDIAEQTNMLA--LNANIEAARAGSGGSGDGFVAVDAVEKELATESORSAXDI 398
QY 280 SN-IDSAXAALTAQTAIEAQAQKPPDPSPILOAEQMVIAQEKDKNKIPADGSDVPVPG 338
Db 399 AELIEVQSQATTVEEI-----ISOAMDEQAQRSERVSS-----VDDIATISQATAD 445
QY 339 TTVGSGKQGGSIGIRVSMMLDDAENETASIMSGFRQMIHMFNTENPDSPQAQOELAA 398
Db 446 ETTDGVQE-----ISOAMDEQAQRSERVSS-----VDDIATISQATAD 484
QY 399 QARAAGDSDSAAALADAQAQALEALGKAGQOQGIIN 437
Db 485 RAENVSAASEEQ-TASITEVTSLSQSLAAQADLTLEDRIN 522
RESULT 12
T46811
halobacterial transducer protein V [imported] - Halobacterium salinarum
C:Species: Halobacterium salinarum
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 15-Sep-2000
C:Accession: T46811
R:Rudolph, J.; Nordmann, B.; Storch, K.F.; Gruenberg, H.; Rodewald, K.; Oesterhelt, D.
FEMS Microbiol. Lett. 139, 161-168, 1996
A:Title: A family of halobacterial transducer proteins.
A:Reference number: Z24094; MUID:96275896; PMID:8674984
A:Accession: T46811
A:Status: preliminary; translated from GB/EMBL/DDJ
A:Molecule type: DNA
A:Residues: 1545 <RUD>
A:Cross-references: EMBL:X95589; NID:g1435:30; PIDN:CAA64842.1; PID:g1435132
C:Genetics: httpv
A:Gene: httpv
C:Superfamily: Halobacterium salinarum transducer protein htrI
Query Match 8.0%; Score 176.5; DB 2; Length 545;
Best Local Similarity 21.4%; Pred. No. 0.029;
Matches 98; Conservative 79; Mismatches 195; Indels 87; Gaps 17;
QY 12 ETERTPPADLSAQGLEASAAKSAEQAQIAGABAKPKESKTSVERWSILRSVNNALMSL 71
Db 118 EEERA-----EAERAREKAQKQAEER--QTAAEASAKQDARERSABIEQLAADLESQ 169
QY 72 ADKLGASSNSSSTSRGADVSTTATPTPPPT-FDDYKTAQATYDTI-----FTST 125
Db 170 ATEVG-ATLEAASDGLTARVDATTDAENAEIAEVAATVNDMLTMTERTIDEIQGFSTNVT 228
QY 126 SLADIQAALVSLQDVTNIKDT---AATDEE---ATAAEWETKNADAVKGAQITE 176
Db 229 ASRETAGAKETODASQTVESVQIEAGTDQREQLSVAEEMDSYATVEEVAATAQS 288
QY 177 LAKYASDNQALDLSLGLT-----SFDLLQAALQSVANNKAAELKEMQDNVPVPGK 230
Db 289 VADTAADTTDVA-TAGQTAEDDAIDAIDAVQETMTTANVDALEDLTTEIDD----- 340
QY 231 TPAIAQSLVDQDTATATQIEKQGNNAIRDYFAGQNASG-----AVENAKNSNSI 279
Db 341 IAEILSDIAEQTNMLA--LNANIEAARAGSGGSGDGFVAVDAVEKELATESORSAXDI 398
QY 280 SN-IDSAXAALTAQTAIEAQAQKPPDPSPILOAEQMVIAQEKDKNKIPADGSDVPVPG 338
Db 399 AELIEVQSQATTVEEI-----ISOAMDEQAQRSERVSS-----VDDIATISQATAD 445

QY 339 TTVGSGKQGGSIGIRVSMMLDDAENETASIMSGFRQMIHMFNTENPDSPQAQOELAA 398
Db 446 ETTDGVQE-----ISOAMDEQAQRSERVSS-----VDDIATISQATAD 484
QY 399 QARAAGDSDSAAALADAQAQALEALGKAGQOQGIIN 437
Db 485 RAENVSAASEEQ-TASITEVTSLSQSLAAQADLTLEDRIN 522
RESULT 13
F84194
HtrI4 transducer [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: F84194
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl, Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L.
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-627 <STO>
A:Cross-references: GB:AB004437; NID:g10579976; PIDN:AAG18922.1; GSPDB:GN00138
C:Genetics:
A:Gene: htrI4
C:Superfamily: Halobacterium salinarum transducer protein htrI
Query Match 7.9%; Score 175; DB 2; Length 627;
Best Local Similarity 22.7%; Pred. No. 0.041;
Matches 116; Conservative 88; Mismatches 190; Indels 116; Gaps 24;
QY 17 PPADLSAQGLEASAAKSAEQAQIAGABAKPK-----ESKTSVERWS-ILRSVNNALMSL 71
Db 145 PALDESVPFAGGESTEMADSLAETAELEKTABLEHQAELEKSQLRALVDALSEA 204
QY 72 ADKLGASSNSSSTSRGADVSTTATPTPPPTFDDYKTAQATYDTIFTSTSLADIQ 131
Db 205 TD-----AARAGLTATVDAAALDVTDDHRAAVERDFNQLLETADTISDIQSFSDAV 256
QY 132 AAL-----VSLQDVTNIKDTA--ATDEETAIAAEWETKNADAVKGAQI 174
Db 257 LAVSRTTDERVDAVADRSAAVSESVEIADGANQTNQNLNIAEEMDTVSATVEIAASA 316
QY 175 TELAKYASDNQAIL---DSLGLK-TSFDLLQAALQSVANNKAAELKEMQDNVPVPG 229
Db 317 NDVAKTA---QAADRGEDGRGEVEETIEALRLREQQVAETVESLAAEVER---IDG 370
QY 230 KTPAIAQSLVDQDTATATQIEKQGNNAIRDYFAGQNASG-----AVENAKNSNS 278
Db 371 IT-ALLIEDIAETNMLAL-----NASIEAARTSGDGFVAVDAVEKDLAETREQAAD 423
QY 279 INIDSA-----KAAIATAKTQIABAQKPPDSP-ILQEAQMVIAQEKDLKNKIPA-- 329
Db 424 ISEIVDAVTEKAEDASIAIGEVD-AEVERKITKAEGVLRDFEAIY---DEVANNHVAQ 478
QY 330 DGSDVPNPG---TTVGSKQGGSSI-----GSRVSMMLDDAEN 365
Db 479 EISDATDQQAQSVTDVVGWVEVASVSEETAESVTADNAAEQTDATDEVADQDELAE 538
QY 366 ETASI--LMSGFRQMIHMFNTENPDSPQAQOELAAQARA--AAGDSDSAAALADAQA 421
Db 539 QTAALAGMLDDF-----TVPADAGTADQSVADDSPTAQPPAADDEPAAAVVDPQFPA 590
QY 422 LEAALGKAGQOQGIINALGQIASAAVVSAG 451
Db 591 SDAE-----DEBGPVDSGGE--SVAVSDGG 613

```

Query Match          7.8%; Score 172.5; DB 2; Length 1128;
Best Local Similarity 23.9%; Pred. No. 0.12;

```

Search completed: March 24, 2004, 06:00:10
Job time : 14.3321 secs

Search completed: March 24, 2004, 06:00:10
Job time : 14.3321 secs

QY 613 FAGSRPTAAEIKALSFET 630
Db 1442 -DELATKANEOKALIAQT 1458

RESULT 8
B90835
probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: B90835
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90835
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035073.1; PID:q13361114; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs1650

Query Match 7.0%; Score 221.5; DB 2; Length 971;
Best Local Similarity 21.5%; Pred. No. 0.002;
Matches 140; Conservative 98; Mismatches 290; Indels 123; Gaps 20;

QY 26 LEASANKSAEAOIRIAGAEAKPKESKTSVERWSILRSVAVNLMSLADKLGIASSNSSS 85
Db 110 VEEVARNASAVAQNTAAAKKSADASTSAREATHATDAADSARAASSTAGQAAS-SAQS 168

QY 86 TSSADVSTTATAPPPPTFDYKTOAQATDTTFTSLADIQAALVSLQDAVNTIK 145
Db 169 ASSAGTASTKATEASKSAASAAESSKSAATSAAGAKTSET-----NAAVSOQSAATSAS 223

QY 146 DTAATDEETAIAEWETKNADAVKGAQITELAKYASDNQAILDSLGKLTSPDLOAALL 205
Db 224 TATTKASEAASSA-----RDASAKSAKSETSAASSASSASATAGNS-----AKAA 274

QY 206 QSVANNKKAELLKEMQCNFVPGKTPAIAQSLVDQDTATQIEKGNIRDAYFAGON 265
Db 275 KTSETNAKSSETAEQSAASAAAGSKTAAALSASAASTSA-----GQASASATAAGKS 326

QY 266 ASGAVENAKNNSISNIDSAKAAIATATQIAEAKKFPDPSPILOAEQVIOAEKDLKN 325
Db 327 A-----ESAASSASTATTKAGEATEQ-----ASAAASSASAATKSETN 364

QY 326 IKPADGSDVNPFGTTVGGSKQGGSSIGSIRVSMLLDDAENETASILMSGFRQIMHFNTE 385
Db 365 AKASSETSAESSKTAASASSASSASSASASK--DEATROASAASATTAATKATEAA 422

QY 386 NPDSQAQOELAAQA-----RAKAGDDSAALADAO----- 419
Db 423 GSATAAAQKSTAESATRAETAAKRAEDIASAVALDEASTTKKGIQVLSATNSTSESL 482

QY 420 ----KALEAALGKAG-----QQQILNALGQIASAAVVSAGVPPA----- 455
Db 483 AATPKAVKAAAYELANGKYTAQDATTAKGIVQLSNATNSTSEMLAATPKSVKAAAYDLANG 542

QY 456 --AASSIGSSVKOLYKTSKSGDYKTOISAGYDAYKSIND-AYGR---ARNDATRDVIN 509
Db 543 KYTAQDATTAKGIVQLSATNSASSET-LAATPKAVKAANDNANGRVPSARKVNGKALSS 601

QY 510 NVS-TP---ALTRSVPRARTARG-----PEKTOALARVIGNSRLTGDVYSQVS 556
Db 602 DITLTPKDIGTLNSTTMSFGGAGWFKLATVTPQASSVSVITLIGGAGFVNGSP-QQAG 660

QY 557 ALQSVNQIIOQNPOANNEIRKQLTSVATKPPQFGYPYVQLSNDSTOKFIA 607
Db 661 ISELVLRAGNGNFKGITGALWQRTSTGFT-----NFAWNTSGDITYIVA 706

RESULT 9
C85693
probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli (stra
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: C85693
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grobbeck, B.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <STO>
A:Cross-references: GB:AE005174; NID:gl2514847; PIDN:AA056007.1; GSPDB:GN00145; UWGP:Z1
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1918

Query Match 7.0%; Score 221.5; DB 2; Length 973;
Best Local Similarity 21.5%; Pred. No. 0.0021;
Matches 140; Conservative 98; Mismatches 290; Indels 123; Gaps 20;

QY 26 LEASANKSAEAOIRIAGAEAKPKESKTSVERWSILRSVAVNLMSLADKLGIASSNSSS 85
Db 112 VEEVARNASAVAQNTAAAKKSADASTSAREATHATDAADSARAASSTAGQAAS-SAQS 170

QY 86 TSSADVSTTATAPPPPTFDYKTOAQATDTTFTSLADIQAALVSLQDAVNTIK 145
Db 171 ASSAGTASTKATEASKSAASAAESSKSAATSAAGAKTSET-----NAAVSOQSAATSAS 225

QY 146 DTAATDEETAIAEWETKNADAVKGAQITELAKYASDNQAILDSLGKLTSPDLOAALL 205
Db 226 TATTKASEAASSA-----RDASAKSAKSETSAASSASSASATAGNS-----AKAA 276

QY 206 QSVANNKKAELLKEMQCNFVPGKTPAIAQSLVDQDTATQIEKGNIRDAYFAGON 265
Db 277 KTSETNAKSSETAEQSAASAAAGSKTAAALSASAASTSA-----GQASASATAAGKS 328

QY 266 ASGAVENAKNNSISNIDSAKAAIATATQIAEAKKFPDPSPILOAEQVIOAEKDLKN 325
Db 329 A-----ESAASSASTATTKAGEATEQ-----ASAAASSASAATKSETN 366

QY 326 IKPADGSDVNPFGTTVGGSKQGGSSIGSIRVSMLLDDAENETASILMSGFRQIMHFNTE 385
Db 367 AKASSETSAESSKTAASASSASSASSASSASASK--DEATROASAASATTAATKATEAA 424

QY 386 NPDSQAQOELAAQA-----RAKAGDDSAALADAO----- 419
Db 425 GSATAAAQKSTAESATRAETAAKRAEDIASAVALDEASTTKKGIQVLSATNSTSESL 484

QY 420 ----KALEAALGKAG-----QQQILNALGQIASAAVVSAGVPPA----- 455
Db 485 AATPKAVKAAAYELANGKYTAQDATTAKGIVQLSNATNSTSEMLAATPKSVKAAAYDLANG 544

QY 456 --AASSIGSSVKOLYKTSKSGDYKTOISAGYDAYKSIND-AYGR---ARNDATRDVIN 509
Db 545 KYTAQDATTAKGIVQLSATNSASSET-LAATPKAVKAANDNANGRVPSARKVNGKALSS 603

QY 510 NVS-TP---ALTRSVPRARTARG-----PEKTOALARVIGNSRLTGDVYSQVS 556
Db 604 DITLTPKDIGTLNSTTMSFGGAGWFKLATVTPQASSVSVITLIGGAGFVNGSP-QQAG 662

QY 557 ALQSVNQIIOQNPOANNEIRKQLTSVATKPPQFGYPYVQLSNDSTOKFIA 607
Db 663 ISELVLRAGNGNFKGITGALWQRTSTGFT-----NFAWNTSGDITYIVA 708

hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
 C:/Species: Staphylococcus aureus
 C:/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:/Accession: B89921
 R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A:/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:/Reference number: B89758; MUID:21311952; PMID:11418146
 A:/Accession: B89921
 A:/Status: preliminary
 A:/Molecule type: DNA
 A:/Residues: 1-6713 <KUR>
 A:/Cross-references: GB:BA000018; PID:gl3701232; PIDN:BA842527.1; GSPDB:GN00149
 A:/Experimental source: strain N315
 C:/Genetics:
 A:/Gene: ebha

Query Match 6.9%; Score 220; DB 2; Length 6713;
 Best Local Similarity 21.6%; Pred. No. 0.03;
 Matches 155; Conservative 106; Mismatches 279; Indels 177; Gaps 32;

QY 27 BASAANKAE-----QRTAGAEAKPKESKTSVERWSILRSVAVNALMSLADKLGIASSNS 82
 DB 2200 KASQNSAKSALNGDEKLAAKQTAKSIGRLTLNNAQRTAANAARVDQAPNLAARVTAAK 2259

QY 83 SSSTSR-----ADVSTTAT-----APTPPTPTFDYKTAQATYDTTFTSTSL 127
 DB 2260 NKATSLNTAMGNLKHAEKWNKESVNYTDADQPKQAYDTAVTQAEAINWANGSNANE 2319

QY 128 ADIQALVSLQDA-----VTNKKDTAATDEETAIAAEWETKNADAVKVGQITELA 178
 DB 2320 TQVQAALNQLNQAQNDLNGDNKVAQAKETA-----KRALASYNLNNAQSTAATSQI----- 2371

QY 179 KYASNQALDLSGLKTSFDLLOALLQSVAN--NNKAELLKEMQDNPVPGKTPATAQ 236
 DB 2372 ----DNAT-----TVADVTAA-----QNTAELNLTAMQLONGINDQNTV-----K 2408

QY 237 SLVDDOTATATQIEKGNRAIRDAYF--AGQNASGAVENAKSNNSISNIDSAKAAIAATAKTQ 295
 DB 2409 QQVNFDDADQK-----KQAYTNVNTAOGILDKANGON-----MTKAQVEAALNQ 2454

QY 296 IAEAQKFPDPILOAEQWJQAEKD-----LNKIKPADGSDVPNPFTTGG-- 343
 DB 2455 VTTAKNAL-----NGDANVRQAQSDAKANLGTTLHLNNAQKODLTSQIEGATTVNGVN 2507

QY 344 ---SKQOG-----SSIGS---IRVSMLLDDAENETASILMSGFQMIHMENTE--- 385
 DB 2508 SVYTKAQDLDGAVQRLSAIANKDQTKASENIDADPTKTKAFDQNAITQAESYLNKHDT 2567

QY 386 NPDQAAQOELAAQRAAKAA--GDDSAALADACKALEAALGKAGQOQOGLNALGQIAS 444
 DB 2568 NKDKQAVEQAIQSVTSTENALNGDANLQCAKTEATQAINLTLQNTPPK---TALKQOVN 2624

QY 445 AAVVSGVPPA--AASIGSSVKQLYKTSKTSQSGDYKTOISAGY-----DAYKSINDAY 496
 DB 2625 AAQRVSGVTDLKNASATSLNNAQDQ-----KQAIQGDHDTIVAGGNYTNASPDQKQAYTDAY 2680

QY 497 GRARN--DATRDVINNV--TPALTQSVPRARTEARG-----PEKTDQALARV----- 540
 DB 2681 NAAKINIVGPNVITNADVTAAITQVRVNNATSLNGDNTNATQAKQAKDALRGVTHLSDA 2740

QY 541 ----ISGNSRTLGDVYSQVSALQSMQIIQS-----NPOANNEIRQKLTSAVTKP 587
 DB 2741 QKQSITQI-----DSATQVTGVQSVKDNAITNLNANMQLNSIANKDEVKA----- 2787

QY 588 PQGYFPVQVQNSDSTQK-----IAKLRLPAEGSRTRAEIKALSPFNLSFIQOVLVN 641
 DB 2788 ---SQPYVDADTKQNAVNTAVTSAENIINATSQPTLDPSAVTQAANQVNTNKTALN 2841

RESULT 11

G64887
 probable tail fiber protein GP37 - Escherichia coli (strain K-12)
 C:/Species: Escherichia coli
 C:/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:/Accession: G64887; T09189
 R/Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:/Title: The complete genome sequence of Escherichia coli K-12.
 A:/Reference number: A64720; MUID:97426617; PMID:9278503
 A:/Accession: G64887
 A:/Status: nucleic acid sequence not shown; translation not shown
 A:/Molecule type: DNA
 A:/Residues: 1-1122 <BLAT>
 A:/Cross-references: GB:AE000234; GB:U00096; NID:91787633; PIDN:AAC74454.1; PID:gl787633
 A:/Experimental source: strain K-12, substrain MG1655
 R/Aiba, H.; Baba, T.; Fujita, K.; Hayashi, K.; Inada, T.; Isono, K.; Itoh, T.; Kasai, H.; Motomura, K.; Nakade, S.; Nakamura, Y.; Nashimoto, H.; Nishio, Y.; Oshima, T.; Saito, Y.; Horiuchi, T.
 DNA Res. 3, 363-377, 1996
 A:/Title: A 570-Kb DNA sequence of the Escherichia coli K-12 genome corresponding to the
 A:/Reference number: Z16603; MUID:97251357; PMID:9097039
 A:/Accession: T09189
 A:/Status: preliminary; translated from GB/EMBL/DBJ
 A:/Molecule type: DNA
 A:/Residues: 3-1122 <AIB>
 A:/Cross-references: GB:AE000234; GB:U00096; NID:91787633; PIDN:AAC74454.1; PID:gl787633

Query Match 6.6%; Score 210; DB 2; Length 1122;
 Best Local Similarity 20.4%; Pred. No. 0.009;
 Matches 154; Conservative 102; Mismatches 294; Indels 206; Gaps 25;

QY 8 GPIDETERTPPA-----DISAQGLEASAANKSAEAORTAGAEAKPKESKTSVERWSILRS 63
 DB 93 GANTEDDARPEALRRFFELM---VEEVARNASAVAQNTAAAKSKSDASTSAREANTHAAD 149

QY 64 AVNALMSLADKLGIASSNSSTSRSDADVSTTATAPTPPTPTFDYKTAQATYDTTFT 123
 DB 150 AADSARAASTSAGQAS--SAQSASSSAGTASTKATEASKSAASAAESSKSAATSAAGAAT 208

QY 124 S--TSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKYA 181
 DB 209 SETNWSASLQSAATSASTATTKASE--AATSARDADAASKEAAKSSE-----TNASSSA 259

QY 182 SDNQALDLSGLKTSFDLLOALLQSVANNKKAELLKEMQDNPVPGKTPATAQSLVDQ 241
 DB 260 SSAASATAAGN-----SAKAATSETNARSSETAAGQASAAAGSKTAAASASAS 312

QY 242 TDATATQIEKGNRAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIAATAKTQIAEAO 301
 DB 313 TSA-----GQASASATAAGKSA-----ESAASASTATTKAGEATE 348

QY 302 KFPDSPILQEBQWVIOAEKLNKIKPADGSDVPNPFTTGGSKQOQSGSISIRVSMLLD 361
 DB 349 Q-----ASAAARSASAAKTSETNAKASE-----TSAESSKTAASASSASSAAS 390

QY 362 DAENETASILMSGFQMIHMENTENFDSQAQOELAAQRAAKAAGDDSAALADACKA 421
 DB 391 ----SAS-----SASAKDEATROASAAKSATTSATKATEAAGSA 427

QY 422 LEAALGKAGQOQOGLNALGQIASAAVVSAGVPPAAASIGSSVKOLYKTSKTSQSGDYK 481
 DB 428 TAAQSKSTAESAATRAETAAKRAEDIAVALEADSTTKKGIVQL---SSATNSTSETL 484

QY 482 ISAGVDYKSIDNAYGRARNDAITRD-----VINNVSTPALT-----RSVPRART 525
 DB 485 AATP-----KAVKSAVDNAEKRLQKQONGADIPDKGCFLLNNINAVSKTDFADKGRMYRV 540

QY 526 EA-----RGPEKTDQALARV--ISGNSRTLGD----- 550
 DB 541 NAPAGATSKYYPVVMRSAGSVSELASRVIITTRATAGDPMNCEFGFVMPGWTDR 600

QY 551 ---VVSQVSALQSMQIIQSNPQAN--NEIRQKLTSAVTKPPQFGYPVYVQLS----- 598


```

A;Gene: SPI772

Query Match          6.5%; Score 208.5; DB 2; Length 4776;
Best Local Similarity 20.3%; Pred.No. 0.071;
Matches 126; Conservative 106; Mismatches 336; Indels 53; Gaps 12;

QY      13  TERTPPADLSAQGLGAEAGAAKNSAEORIAQAEAKPKESKTDTSVERHSILRSAYNALMSLA 72
DB      2283 TSASVSASTSASASASTSASESASTSASASASTSASESASTSASASTSASASTSASASASTS-A 2341

QY      73  DKLGIASSNSSSTRSADVDTATPTPPPPFDDYKTAQAVDTITTSGLADIQA 132
DB      2342 SESASTSASASTSASASA-STASA-----SASTSASESASTSAS---A 2385

QY      133 ALVSLQAVTNIKDTAATD-----EETAIAABWETKNADAVKVGCAQTTELAKVASDNQAI 187
DB      2386 STSASEASTSASASTSASASTSASASTSASASTSASASTSASASTSASESASTSASASTSASA 2445

QY      188 LDSLGKLTSTFDLLOAQLQSVAANNKAAELLKEMQDNPPVPGKTPPAIAQSLVDTQDATAT 247
DB      2446 SASTSASASASTSASASTSASASTSASASTSASASTSASASTSASASTSASASTSASAST 2505

QY      248 QIEKDGNAIRDAYPAGQNAS-----GAVENAKSNNSISNIDSAKAAATATAKTOIAEAOQK 302
DB      2506 SASASASTSASA-SASTSASASTSASESASTSASASTSASASTSASASTSASASTSASASTSAS 2564

QY      303 FPDSPILLQAEQMWIQAEKOLNIKPKADGSDVPNPGTTVGGS-KQCGSSIGCSIRVSMLLD 361
DB      2565 VSASTSASESASTSASASTSASASTSASASTSASASTSASESASTSASESASTSASASTSASASAS 2624

QY      362 DAENETASILMSGPROMHMENTENPQSAQQELAAQARAAXKAAGDDSAALADQKA 421
DB      2625 TSASEASTSASA-----SASTSASASTSASASTSASTSASASTSASASTSASASTS 2676

QY      422 LEMALGRAGQQQGIILNALGQIASAAVVSAGVPPPAASISGSSVKQLYKTSKSTGSDYKTO 481
DB      2677 ASAGASTSASASASTSASESASTSASASTSASASTSASASTSASASTSASESASTSASASTSASAS 2736

QY      482 ISAGYDAYKSYIN-DAYGRARNDAFRDVINNVSTPALTRSVPRAPTEARGPEKTDQALARY 540

```

Db 2737 EGASTSASASTSTGASASTSASESASTSASASASTSASASASTSASASASTSASASASTS 2796
QY 541 IGSNSRTLGDVYSQVSAQVSMQIIQSNQANNEIRQKLTSAVTKPPQFGYVYVQLSND 600
Db 2797 ASESASTSASASTSA--SASASTSASASTSASASASTSASASTSASASTSASASTS 2841
QY 601 STQKFIKLESLEFAESRTAA 621
Db 2842 STSASVSASTSASASASTSAS 2862
RESULT 14
T30822
lmp1 protein - Mycoplasma hominis
C:Species: Mycoplasma hominis
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 07-Dec-1999
C:Accession: T30822
R;Jensen, L.T.; Ladefoged, S.; Birkelund, S.; Christiansen, G.
Infect. Immun. 63, 3336-3347, 1995
A:Title: Selection of Mycoplasma hominis PG21 deletion mutants by cultivation in the pre
A:Reference number: Z18884; MUID:95369882; PMID:7543881
A:Accession: T30822
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1365 <JEN>
A:Cross-references: EMBL:U21962; NID:g790243; PID:g790244; PIDN:AAA81013.1
C:Genetics:
A:Gene: lmp1
A:Genetic code: SGC3
Query Match 6.2%; Score 199; DB 2; Length 1365;
Best Local Similarity 19.1%; Pred. No. 0.04; Mismatches 276; Indels 202; Gaps 28;
Matches 143; Conservative 126;
QY 13 TERPPADLGAQGEASAANKSAAQRIAGAEAKPKESKTDVSVERWSILRSVAVNALMSLA 72
Db 317 TDSNKSDESANTELKQALAKANADKV-----QADNLAKSIK 354
QY 73 DKLGIASNSSSSTSRADVDSTTATATPPPTFDYKTAQAYDTITFTSLADIQA 132
Db 355 EQLNNSVSNANTLSAKLTDKNTIQAQK-----ELEKEVQKA-DQAIKSNNTASMQS 406
QY 133 ALVSLQDAVNIK--DTAATDETAATAAEWETKNADAVKGAQ-----ITELAKY 180
Db 407 AKSLDQKVAEITKKLFTFNKQKAEKFNELKQTRNQIQEFINTKNPNYSSELLISQTSK 466
QY 181 ASDNQAILDSGLKLTSPDLQAALLQSVANN-----KAAELLKEMQDNPPVPGKTPA 233
Db 467 RDSKNSVTDSSNK--SDIESANTELKQALAKANADKVQADNLAKSIKEQLNNSVSNANT-- 523
QY 234 IAQSLVDTDA---TATQIEKQGNARIDAYPAGQVAGAVENAKSNNSISNIDSAKAA 290
Db 524 LSAKLTDKNTIQAQKTELEKE-----VQKADQAI---KSNNTAS--MQSAKSSL- 568
QY 291 TAKTQIAEAKKFPDPSPIQAEQVMIQAEKDLKNIKPADGSDVNPPTTGGSKQGGSS 350
Db 569 --DAKVAETTKL--ETFNKQKAEKFNELKQTRNQIQEFINTKNPNY----- 613
QY 351 IGSIRVSMLLDDAENETASILMSGFQMIHMFNTENPDQAAQELAAQAPAAAGDSDS 410
Db 614 --SELISQLTSKRSKN--SVTDSNKSNDIESANTE-----LQALAKANADKVQADNL 663
QY 411 A-----AAALAD-----AQKALEAALGKAGQOQILNALQIASAA 446
Db 664 AKSIEQLNNSVSNANTLSAKLTDKNTIQAQKTELEKEIQKANO-----AIKSNNTAS 717
QY 447 VVSAGVPAAPAASSIGSSVKQLYKTSKSTGSDYKTIQISAGYDAYKSIDNAYGRAND---- 502
Db 718 MQS-----AKSLDAKVAETTKLFTFNKQKAEKFNELKQTRNQIQEFINTKNPNYS 771
QY 503 -----ATRDVINNVSTPALTRSPVPRARTFARGPEKTDQALARVISGNSRTL-GDYVSQ 554
Db 772 ELISQLTSKRSKNSVTDSSNKSNDIESANTELKQALNTAKAKKSSIDNELPLKNDLQSK 831

QY 555 V-----SALQSV-----MQIIQSNQANNEIR-----OKLTSA 583
Db 832 IBEFGPIRINTNFSWISSKLETTKKNLABELTKADAIKNNPSSKQALXDSQOQVKLGNE 891
QY 584 V-----TKPPQFGYVYVQLS-----DSTQKFIKLESLEFAE 615
Db 892 LLKTYTEBFGKVETKNSNIGYRLFKLQAQEPNNSDVVKLKNWEEKQTLSSKKQKLGQ 951
QY 616 GSRTAAEIKALSFETNSLFIQQVLVNI 642
Db 952 STKDYLTQLSTEMSTQESTIKKIVNI 978
RESULT 15
T24583
hypothetical protein T06D8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24583
R;Palmer, S.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z19909
A:Accession: T24583
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1829 <WIL>
A:Cross-references: EMBL:Z49130; PIDN:CAA88964.1; GSPDB:GN00020; CESP:T06D8.1
C:Experimental source: clone T06D8
C:Genetics:
A:Gene: CESP:T06D8.1
A:Map position: 2
A:Introns: 1391/3; 1432/3; 1470/3; 1505/1; 1520/1; 1616/1; 1644/1; 1687/3; 1742/1
Query Match 6.2%; Score 196.5; DB 2; Length 1829;
Best Local Similarity 20.5%; Pred. No. 0.078;
Matches 144; Conservative 108; Mismatches 303; Indels 147; Gaps 26;
QY 18 PADLSA--QGLEASAANKSAAQRIAGAEAKPKESKTDVSVERWSILRSVAVNALMSLADKL 75
Db 68 FTLSKDDQVTEAS---GEETTAATAEASSEETTSATVTEGSGEETTVVAVVESGGE- 122
QY 76 GIASSNSSSSTSRADVDSTTATATPPPTFDYKTAQAYDTITFTSLADIQAALV 135
Db 123 EPASSSTSVFTLSKDDQVTEASG-----EETTTAAATEASSEETTSATVTEGSGEE 174
QY 136 SLDQAVTNIKDTAAT-----DEETAAAEWETKNADAVKGAQ 174
Db 175 TTTSAVTEASSEATTTTTPAGTEASGEETTSATVTEGSGEETTVVAVVESGGEPPASSSTSI 234
QY 175 -TELAKYASDNQAILDSGLKLTSPDLQAALLQSVANNKAAELLKEMQDNPPV-----P 228
Db 235 PTLSK-----NDQVTEASGEET---ITAAATEASSETTSATVTEGSGEDTTVVAVVELS 286
QY 229 GKTPALIAQSLVDTDATATQIEKQGNARIDAYPAGQVAGAVENAKSNNSISNIDSAKAA 288
Db 287 GEQPASSSTSI-----PTLSKD-DQVTEA--SGEETTTAAATEASSEETTSATVTEGSG 337
QY 289 IATAQIABAQKFFPDS-----PIQAEQVMIQAE-----EKDLKNIKP 328
Db 338 EETTVVAVVESGGEPPASSSTSIPTLSKDDQVTEASGEETTTAAATEASEETTTTSAVTE 337
QY 329 ADGSDVNPPTTGGSKQGGSS--GSIKRVSMLLDD-----AENETASI 370
Db 398 GSGEDTTVVAVVESGGEQPASSSTSIPTLSKDDQVTEASGEETTTAAATEASEETTTSA 457
QY 371 LMSGFQ-----MIHMFNTENPDQAAQ--QELAAQAPAAAGDSDSAAAAALADA--QKA 421
Db 458 VTEGSGEDTTVVAVVESGGEQPASSSTSIPTLSKDDQVTEASGEETTTAAATEASEETTT 517
QY 422 LEAALGKAGQOQILNALQIASAAVVSAGVPPAAAS--SIGSSVKQLYKTSKSTGSDYKT 480
Db 518 TSAVTEGSGEETTV-----AVVESGGEPPASSSTSIPTLSKDDKDVTEASGEETTT 569

Qy	481	QISAGYDAYKSINDAYGRANDATRDVI-----NNVSTPALTTRSVPRAR-----TEARGP	530
Db	570	AAATDASSEETTTSAVTEGSGEETTVAVVRESSDEEPASSSTSIPTELSKDDQVTEASGE	629
Qy	531	EKTDQALARVISGNSRTLGDVYSQVSALQSVMOIIQSNPQANNEEIRQKLTSAVTKPPQF	590
Db	630	ETTTAAATE--ASEETTTSAVTEGSGEETTVAVVES-----SGEEPASSSTSIPT-----	678
Qy	591	GYPVQLSNDSTQKFIAKLESIFAEGSGRTAAEIKALSFEYNS	632
Db	679	-----ELSKDD-----KVTEASGEETTTAAATDASSEETTT	709

Search completed: March 24, 2004, 06:00:03
Job time : 28.7616 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 19:25:30 ; Search time 17.7616 Seconds
(without alignments)

1892.204 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MNPPIGPGIDETERTPPAD.....SLFTQQVLVNIIGSLYSGYLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/6C_COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/6D_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2182	68.5	478	4	US-09-198-452A-776
2	1000	31.4	212	4	US-09-198-452A-775
3	671.5	21.1	361	4	US-09-556-877-299
4	671.5	21.1	361	4	US-09-620-412C-299
5	671.5	21.1	361	4	US-09-598-419-299
6	214.5	6.7	2137	4	US-09-134-001C-4463
7	210	6.6	10182	4	US-09-134-001C-3159
8	186.5	5.9	3696	4	US-09-134-001C-5080
9	178.5	5.6	8991	4	US-08-714-741-32
10	173.5	5.4	1561	3	US-08-894-017-23
11	173.5	5.4	1561	4	US-09-456-474-23
12	173	5.4	2285	4	US-09-308-375-2
13	170.5	5.3	718	4	US-09-540-236-2753
14	167	5.2	1566	2	US-08-687-956A-23
15	165.5	5.2	2310	4	US-08-874-923-120
16	165	5.2	1565	6	535250-2
17	164.5	5.2	469	4	US-09-489-039A-13565
18	164	5.1	1129	4	US-09-252-991A-29927
19	160.5	5.0	639	1	US-08-466-390-2
20	160.5	5.0	639	1	US-08-470-950-2
21	160.5	5.0	639	1	US-08-467-781-2
22	160.5	5.0	639	1	US-08-195-487-2
23	160.5	5.0	639	1	US-08-483-924-2
24	160.5	5.0	639	5	PCT-US93-06160-2
25	158.5	5.0	468	4	US-09-328-352-6321
26	157.5	4.9	1093	5	PCT-US93-03077-1
27	157.5	4.9	1848	3	US-08-296-791-6

ALIGNMENTS

RESULT 1
US-09-198-452A-776
; Sequence 776, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention of infection and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 776
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-776

Query Match 68.5%, Score 2182; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.5e-144;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MNPPIGPGIDETERTPPADLSAAGLEASAAKSAEQAQRIAGAEAKPKESKTDVSVERWSI	60
Db	3	MNPPIGPGIDETERTPPADLSAAGLEASAAKSAEQAQRIAGAEAKPKESKTDVSVERWSI	62
Qy	61	LRSAVNALMSLADKLGIASSNSSSTSRSDVDSVTTATPTPPPTDDYKTAQTAYDT	120
Db	63	LRSAVNALMSLADKLGIASSNSSSTSRSDVDSVTTATPTPPPTDDYKTAQTAYDT	122
Qy	121	IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKGAQITELAKY	180
Db	123	IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKGAQITELAKY	182
Qy	181	ASDNOAILDSIGKLTSPDLLQAALLQSVANNKAAELLKEMQDNFVPGKTPAQAQSLVD	240
Db	183	ASDNOAILDSIGKLTSPDLLQAALLQSVANNKAAELLKEMQDNFVPGKTPAQAQSLVD	242
Qy	241	QTDATATQIEKGNAINRDYFAGNAGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ	300
Db	243	QTDATATQIEKGNAINRDYFAGNAGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ	302
Qy	301	KKFPDPSILOAEQMVIOAEKDKLNKIKPADGSDVDPNPGTTVGGSKOQSGSISGRVSMLL	360
Db	303	KKFPDPSILOAEQMVIOAEKDKLNKIKPADGSDVDPNPGTTVGGSKOQSGSISGRVSMLL	362
Qy	361	DDAENETASIIIMSGFRQMIHMFENTENPDSQAQQLAAQAAKAAAGDSDSAAALADAQK	420
Db	363	DDAENETASIIIMSGFRQMIHMFENTENPDSQAQQLAAQAAKAAAGDSDSAAALADAQK	422

QY 421 ALRAALGKAGQQQGIINLALQIASA 447
 Db 423 ALRAALGKAGQQQGIINLALQIASA 449

RESULT 2
 US-09-198-452A-775
 ; Sequence 775, Application US/09198452A
 ; Patent No. 6559294
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffiths, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/09/198,452A
 ; CURRENT FILING DATE: 1998-11-24
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 775
 ; LENGTH: 212
 ; TYPE: PR1
 ; ORGANISM: Chlamydia pneumoniae
 US-09-198-452A-775

Query Match 31.4%; Score 1000; DB 4; Length 212;
 Best Local Similarity 99.5%; Pred. No. 2.9e-62;
 Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 448 VSAGVPPAAASSGSSVKQYKTSKSTGSDYKQISAGYDAYKSIINDAYGRANDATRDV 507
 Db 9 VSAGVPPAAASSGSSVNQYKTSKSTGSDYKQISAGYDAYKSIINDAYGRANDATRDV 68

QY 508 INNVTALTRSPVPRTEARGPEKTDQALARVISGNSRSLGVDYQVSVLQSVNQIQQS 567
 Db 69 INNVTALTRSPVPRTEARGPEKTDQALARVISGNSRSLGVDYQVSVLQSVNQIQQS 128

QY 568 NPQANNEEIRKLTSAVTKPQPGYVQVLSNDSTQKFIKLSFAEGSRRTAAEIKALS 627
 Db 129 NPQANNEEIRKLTSAVTKPQPGYVQVLSNDSTQKFIKLSFAEGSRRTAAEIKALS 188

QY 628 FETNSLFIQQLVNLGSLYSGYLQ 651
 Db 189 FETNSLFIQQLVNLGSLYSGYLQ 212

RESULT 3
 US-09-556-877-299
 ; Sequence 299, Application US/09556877
 ; Patent No. 6432916
 ; GENERAL INFORMATION:
 ; APPLICANT: Probst, Peter
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir
 ; APPLICANT: Fling, Steve
 ; APPLICANT: Maisonneuve, Jeff
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.469C5
 ; CURRENT APPLICATION NUMBER: US/09/556,877
 ; CURRENT FILING DATE: 2000-04-19
 ; NUMBER OF SEQ ID NOS: 305
 ; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
 ; SEQ ID NO 299
 ; LENGTH: 361
 ; TYPE: PR1
 ; ORGANISM: Chlamydia
 US-09-556-877-299

Query Match 21.1%; Score 671.5; DB 4; Length 361;
 Best Local Similarity 42.4%; Pred. No. 4.5e-39;
 Matches 154; Conservative 62; Mismatches 130; Indels 17; Gaps 8;

QY 301 KXFPDSPILQAEQWVIOAEKDKNIKPADGSDVPNPGTTVGSKQQGSSIGSI----- 354
 Db 2 QEADSPLVKAEQINQAOQDIQTTPS-GLDIPVGPS-GSAASAGSAGALKSSNNS 59

QY 355 -RVSMILLDDAENETASILMSGFRMHMENTENPDSQAQOELAAQARA---AAGDD 410
 Db 60 GRISLLDDVDNEMAAIAMQGFMSIEQFNVNPNATAKELQAMEAQLTAMSDQLVGADGE 119

QY 411 AAAALADAKALEAALGKAGQQQGIINLALQIASA 449
 Db 120 LPAEIQAIKDALAAL-KOPSADGLATANGQVAFAAKVGGSAGTAGTVMNVKLYKT 178

QY 471 --SKTSGDYKTIQISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPVPRTEAR 528
 Db 179 AFSSTSSSSYAAALSDDYSAYKTLNLSYSESR-GVQSAISQTNPALSRVSRSRGIESQ 237

QY 529 G-PEKTDQALARVISGNSRSLGVDYQVSVLQSVNQIQQS 567
 Db 238 GRSADASQRAETIIVRDSQTLGVDYVSRLOVLSLSTIVSNPQANQEEIMQKLITASISKA 297

QY 588 PQFGYPYVQLSNDSTQKFIKLSFAEGSRRTAAEIKALS 627
 Db 298 PQFGYPYVQLSNDSTQKFIKLSFAEGSRRTAAEIKALS 357

QY 648 GYL 550
 Db 358 GYL 360

RESULT 4
 US-09-620-412C-299
 ; Sequence 299, Application US/09620412C
 ; Patent No. 6448234
 ; GENERAL INFORMATION:
 ; APPLICANT: Steven P. Fling
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.469C7
 ; CURRENT APPLICATION NUMBER: US/09/620,412C
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 363
 ; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
 ; SEQ ID NO 299
 ; LENGTH: 361
 ; TYPE: PR1
 ; ORGANISM: Chlamydia
 US-09-620-412C-299

Query Match 21.1%; Score 671.5; DB 4; Length 361;
 Best Local Similarity 42.4%; Pred. No. 4.5e-39;
 Matches 154; Conservative 62; Mismatches 130; Indels 17; Gaps 8;

QY 301 KXFPDSPILQAEQWVIOAEKDKNIKPADGSDVPNPGTTVGSKQQGSSIGSI----- 354
 Db 2 QEADSPLVKAEQINQAOQDIQTTPS-GLDIPVGPS-GSAASAGSAGALKSSNNS 59

QY 355 -RVSMILLDDAENETASILMSGFRMHMENTENPDSQAQOELAAQARA---AAGDD 410
 Db 60 GRISLLDDVDNEMAAIAMQGFMSIEQFNVNPNATAKELQAMEAQLTAMSDQLVGADGE 119

QY 411 AAAALADAKALEAALGKAGQQQGIINLALQIASA 449
 Db 120 LPAEIQAIKDALAAL-KOPSADGLATANGQVAFAAKVGGSAGTAGTVMNVKLYKT 178

QY 471 --SKTSGDYKTIQISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPVPRTEAR 528
 Db 179 AFSSTSSSSYAAALSDDYSAYKTLNLSYSESR-GVQSAISQTNPALSRVSRSRGIESQ 237

QY 529 G-PEKTDQALARVISGNSRSLGVDYQVSVLQSVNQIQQS 567
 Db 238 GRSADASQRAETIIVRDSQTLGVDYVSRLOVLSLSTIVSNPQANQEEIMQKLITASISKA 297

QY 588 PQFGYPYVQLSNDSTQKFIKLSFAEGSRRTAAEIKALS 627

Db 298 P Q G G P A V Q N S D S I K Q F A A Q L E R E F V D G E R S L A E S Q E N A F R K Q P A F I Q Q L V N I A S L F S 357
648 GYL 650
Db 358 GYL 360

```

RESULT 5
US-09-598-419-299
; Sequence 299, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-299

```

RESULT 6
 US-09-134-001C-4463
 ; Sequence 4463, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064, 964

[illegible]

```

; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match      6.6%; Score 210; DB 4; Length 10182;
Best Local Similarity 20.4%; Pred. No. 6.4e-05;
Matches 149; Conservative 108; Mismatches 290; Indels 184; Gaps 30;

QY 16 TPADLSAQGLESAANKSBAEAGIAEAE-----KPKES-----KTDSV 55
Db 2628 TPQVNSALSQVAAQKINEAKALLQNKADNSQLVRAKEQLQOOSIQPAASTGMDTQDST 2687
QY 56 ERWSILRSVAVNMLSLADKLGIASSNSSTSRSDVSTTATPTPPPTFFDY-----110
Db 2688 RNYKVRQAAEQAIQHANSV---INNGDATSQQINDAKNIVEQAAQOR-----DYVEAKS 2737
QY 111 -----KTQAQATYDTTFTSTSLADIAQALV-SLQDAVNTNFKDTAATDEETAIAAEWETKN 164
Db 2738 NLRADKSQLQSAVDLNRDLVTKKPAVRRYNEAISNR-----KE 2780
QY 165 ADAVKVGAQITELAKIYASDQAILDSLGLKLTSPD--LLQA-ALLOSVAANNKAAELLKEM 221
Db 2781 LDTAKADASST-LRNTNPSVEQVRDALNKINTVQPKVQAIALLQPKENNELVQAKKRL 2839
QY 222 QDN-----PVVPGKTP-----ALAQSLVQDTATATQIEKGNNAIRDAY 260
Db 2840 QDAVNDIPQTCGTQQTQNTNNVNDKQREARALTSQARVINDGDTATTCQETSEKSKVEQAM 2899
QY 261 FAGONASG-----AVENAKSN-----SISNIDSAAIAIAKTQIABAQK 301
Db 2900 QALTNAKSNLRADKVELQATYKNLIENVSTNGKPKASIRQVETAKARI---QNDKNDKN 2956
QY 302 KFPDSPILQEAQVQVCAEAKDLKNIKPADGSDVNPVPGTGVGSKQOQSSIGSIRVSLLD 361
Db 2957 E--AERILGNDNPQVSVQVQALNKIKAIQ-----PKLTEINMLQKNNTELNVAKNRLE 3010
QY 362 DAENETASILMSGPQMIHMFNTENPDSQAQOELAAQARAQAAAGDSDSAAALADAKA 421
Db 3011 NAVNDTDP--THGWTQ--ETINNNYNAKREAQNEI--QKANMIINNGDATAQDTSSEKSK 3064
QY 422 LEAALGKAGQOQGLNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSTG-----475
Db 3065 VEQVL-----QALQNAKNDLR-----ADKRELOTAYNKLIQNVNTNGKPKSSI 3107
QY 476 SDYKTIQISAGYDAVKSINDAYGRARND-----TRDVIN-----NVSTELTRSV 520
Db 3108 QNYKS-----ARRNTENQYNTAKNEAHNVLENTNPTVNAVEALRKINAIQEVTKAI 3160
QY 521 PRAKTE-----ARGPEKTDQALARVIGNSRTIGDVYSQVSALQSVMLIQS-----N 568
Db 3161 NILQDKEDNSLVRAKEKLDQAINSQPSLNGMTQESINNYTTKREACNIASSADTIINN 3220
QY 569 PQANNEETR-----QKLTSAVTYPPQPGYFYVOLSNDSSTOKFTAKULESFAEGSRFAE 622
Db 3221 GDASIEQITENKIRVEEATNALNEAKQ-----HLTADTTSLKTEVRKLSRRGDTNKKPSS 3276
QY 623 IKALSFFETNSL 633
Db 3277 VSAYNNTIHSL 3287

RESULT 8
US-09-134-001C-5080
; Sequence 5080, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5080
; LENGTH: 3696
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5080

Query Match      5.9%; Score 186.5; DB 4; Length 3696;
Best Local Similarity 21.7%; Pred. No. 0.00067;
Matches 171; Conservative 118; Mismatches 293; Indels 205; Gaps 40;

QY 12 ETERTPADLSAQGLESAANKSBAEAGRI-----AGAEAKPKESKTDSTVERWSILR---- 62
Db 1956 QINNTTDA-TEEBEQEATNRVNAGLAQAIQNTNNNAHSTQEVNESKTNSTATTKSVQPNVI 2014
QY 63 ---SAYNALMSLADKLGIASSNSSTSRSDVSTTATA-----PTPPPTFFD 109
Db 2015 KPETA-NSLTQEBANNOKTLIGNDGNATDDKEKAQKLVTKLNEQIQKHESQODNOVDN 2074
QY 110 YKTAQATA-----YDTIFTSTSLA-----DIOAALVSLQDAVNTNFKDTA--ATD 151
Db 2075 VKAQAITAIKLINANAHRQDAINILTNLAESKSDIRAN---QDATTEKNTAIQSID 2130
QY 152 EETA-----IAEWETK---NAD-AVKVGAQITELAK 179
Db 2131 DTLQAQRNNINGANTVALVDENLEDQKQLRVLSTQTKTQAKADIQAIGQORSTIDQ 2190
QY 180 ---YASDQAIILDSGLKLTSS--FDLLQALL--QSVAN--NNKAAELLKEMQDNVNV---P 228
Db 2191 NQWATTEEQEALERLNQETNGVNDRIQAALANQNVTDKNNILETIRNVE--PIVIVKP 2248
QY 229 GKTPTAQSILVDQDTATATQIEKGNNAIRDAYPAGNAGAVENAKSN--NSISNTDSAK 286
Db 2249 KANEIIRKAAEQT---TLNQNDATLE---EKQIALGKLEVEKNEALNQVSAHSNN 2301
QY 287 AATATAKTQIAEAKKFPDSPILQEAQVQVQ-----AEKDLKN 325
Db 2302 DVKIAENNGIAKISEVHPETIIRKNAKQIEIQDAQSQIDITINANNKSTNEEKSAAIDRVN 2361
QY 326 IKPADG-SDVNPNGTT--VGGSKQOQ--SSIGSI-----RVSMLLDDAENETASILM 372
Db 2362 VAKIDAINNTITNATTTQLVNDKAKNSGNTSISQILPSTAVKTNALAAALAEAKKNKNAII-- 2419
QY 373 SGFRQMIHMFNTENPDSQAQOQELA-----NQARAQAAGDSDSAAALADAKAL 422
Db 2420 ---DQTPNATAEEKEEANKNVDRLOQEBADANILKAHTTDEVNNIKNQAVQNI 2468
QY 423 EAALGKAGQOQGLNALGQ-IASAAVVSAGVPPAAASIGSSVKQLYKTSKSTGSDYKTQ 481
Db 2469 NAYQVEVIKQNVKNQNLQFIDNQKLIENTPDTLEEAEANRLQLNVLTSTSDSI--- 2525
QY 482 ISAGYDAYKSINDAYGRARN-----DATRDVNNVSTP--ALTRSVPRARTBARGP 530
Db 2526 --ANVDHNEVDQALDKARPKIEEIVPQVSKRDLVNAIQEAQFNSQTOEI-QENQEAATNE 2582
QY 531 EKTQ-----QAL--ARVLSGNSRTIGDVYS-QVSALQSVMLIQSNPQAN-----NE 574
Db 2583 EKTEALNKINQLNQAQKVIDQAKNSKNDVDSAKTRSIQIDQEQ-IQHPQTKATGRHLNE 2641
QY 575 EIRQKLTSAVTKPPQPGYFYVOLSNDSSTOKFTAKULESFAEGSRFAEIKALSFFETNSLF 634
Db 2642 KANQOQSTIATHP-----NSTIEERQEAQKLEVLK-----AIKIDKQND-D 2687
QY 635 IQOVLVN 641
Db 2688 VEKTVVN 2694
```



```

;      TOPOLOGY: linear
;      MOLECULE TYPE: pro
US-08-894-017-23

```

Query Match	5.4%; Score 173.5; DB 3; Length 1561;
Best local Similarity	19.9%; Pred. No. 0.0016;
Matches	140; Conservative 118; Mismatches 304; Indels 143; Gaps 27;
QY	48 KESKTDVSERWSILRSVAVNALMSLADKLGIIASSN-SSSSTSRSDVDSTATATATPPPT 106
Db	2 KVKKTYGERKSKISKTLGGAVLGTVAASVAGQKVFADTTTTSDVDTKVVGTOIGNPAT 61
QY	107 -----FDDYKTKQAQYATYDTIFST-----SLADIQAAALYS-LQDAVTNIKDT 147
Db	62 NLPEAQGSASQAQSCQTKLEROMVHTIEVPKTDLQAAKDAKSAGVNVVQADVN-KGT 120
QY	148 AATDEETAAAEWE-----TKNDAVK-----VGAQTTELAKVASDNQAILDSLIGK 193
Db	121 VKTAAE-AVQKETEIKEDYTKQABDIKKTTPDYKSDVAAHAERAVAKIKAKNQATKEQYVK 179
QY	194 LTSFDLI-QAALLQSVANNKAAEL-----LKEQDNPVVPKTPATQAQSLVDQT-----D 243
Db	180 ----DMVAHKKEVERKINANAASKTAYEAKUQYQADLAAVQKTNAAQAGYQKALAAAYQ 235
QY	244 ATATQIEKDGNNAIDFAPAGQNASGAVENAKSNSISNIDSAKAAIATAKTQIAEAKCKF 303
Db	236 AELKRVQEAANAAKAAY-----DTAVAANNAKNTELAAANBEIRKKNATAKA----- 282
QY	304 PDSPIQEAQEMVTOAEKDLKNIHPADGSDVPNPGTTVGG-----SKQGSISISIR 355
Db	283 -----EYETKLAQYQAEKLRVQEAANAANADYCAKLTAYQTELARVKANADAKAAYE 335
QY	356 VSMLLDAAENETASILMSGFPROMHMFENTENPDSQAQOELAAQARAARAAGDSSAAAL 415
Db	336 AAVAANNAKNAALTAENATAIKQ-----RRENAKATYEAALKQTEADLAAVKKANAAE 388
QY	416 ADAQKALEALGKAGQOQGIILNALGQIA-SAAVVSAGVPPAAAASSIGSSVKQ----- 466
Db	389 ADYQAKLTAYQTELARVQK-ANADAKAAYEAAVAANNAANALTAENATAIKRKNADAKAD 447
QY	467 -----LYKTSKSTGSDYKTKQISAGVDYKGINDAYGRARNDATRVVINNVSTPA 515
Db	448 YEAKLAKYQADLAKYQKDL-ADYPVKLKAYEDEQASIKAAALAELEKHKXND--GNLTSPS 504
QY	516 LTRSVPRARTBARGPEKTDQALARVIGSNRTLGDVYSQVAL-----QSVNQ----- 563
Db	505 AQNLIVDYLEPRNATSLATTDGKFLK-----ASAVDDAFSKSTSKAKYDQKIIQLDDELDTN 559
QY	564 IIQGNPOANNEEI-----RQKLTSAVTKPQFGYPYVQLS-NDSQKFIKLESLFAEG 616
Db	560 LEQSDNVASSNELVGNFGDKAGWSITVSNNSQVWGSVLLBERGQSATATYTNLQNSYNG 619
QY	617 SRTAAEIKALSFETNS-----LFIQQVLVNIQSLSYGYLQ 651
Db	620 KKISKIVKYVDPKSKFOGOKWVLGIETDPTLGVFAFYAAYTGOVE 664

```

RESULT 11
US-09-456-474-23
; Sequence 23, Application US/09456474
; Patent No. 6500433
; GENERAL INFORMATION:
; APPLICANT: Lehner, Thomas
; APPLICANT: Kelly, Charles
; TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE OF
; TITLE OF INVENTION: COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/II
; FILE REFERENCE: 25150-20067.10
; CURRENT APPLICATION NUMBER: US/09/456,474
; CURRENT FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: US 08/894,017
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 23
;
; LENGTH: 1561
; TYPE: prt
; ORGANISM: Streptococcus mutans
US-09-456-474-23

```

Query Match	5.4%; Score 173.5; DB 4; Length 1561;
Best Local Similarity	19.9%; Pred. No. 0.0016;
Matches	140; Conservative 118; Mismatches 304; Indels 143; Gaps 27
QY	48 KESKSDSVRSILRSVAVNLSMADLKLGIASSN-SSSSTRSRADVSTSTTATPTPPPT 106
Db	2 KVKTYTGFRKSKLTKLGGVLTGVAASVAGQVFADETTTSDVDTKVGTQTGNPAT 61
QY	107 -----FDDYKTAQATAYDTITFTST-----SLADIOQALIVS-LQDAVNIKDT 147
Db	62 NLPEAQGSASQAEQSQTCLERQMWHTEVPKTDLQAAKDAKSAGNVVQADVN-KGT 120
QY	148 AATDDETAIAEWE-----TKADAVK-----VGAQITELAKYSDNQAILDSLKG 193
Db	121 VKTAE-AYOKETIKEDYTKQAEIDIKTTDOYKSDVAHEAEVAKIKAKNQATKEQYK 179
QY	194 LTFSELL-QAALLQSVANNKAAEL-----LKEMQDNPPVPGKTPAIAQSLVDOT----D 243
Db	180 ----DMVAHKAEVERKINNAASKTAYEAKLAQVQADLAAVQKTNANQASYOKALAAVQ 235
QY	244 ATATQTEKDGNAIRDAYFAGNAGSAGVENAKSNNSISNIDSAKAAIATATQTAEAKQKF 303
Db	236 AELKRVQEAANAAKAAY----DTAAVANNAKNTEIAAANEEIRKNATAKA-----282
QY	304 PDSPIQZABQMVQAEKDLNRIKPDGSDVPNPGTTVGG-----SKQOGSSISGSTR 355
Db	283 -----EYETKLAQYQAEKRVQEAANAADYQAKLTAYQTELARVQKANADAKAAYE 335
QY	356 VSMLLDDAENETASTILMSGFRQMLHMENTNPDSQAQQELAAQAAKAGDSDASAAAL 415
Db	336 AAVANNAKNAALTAENTAIKQ-----RNNKATYEAAALKQYEAADLAAYVKANAANE 388
QY	416 ADAQKALEAALKAGCQOQGIINALGQIA-SAAVVSAGVPPAPAAASSIGSSVKQ-----466
Db	389 ADYQAKLTAYQTELARVQK-ANADAKAAYEAAVAANNAANAALTAENTAIAKKENADAKAD 447
QY	467 -----LYKTSKSTGSDYKQTSISAGVDAYKSINDAYGARNDATRDVINNVSTPA 515
Db	448 YEAKLAQYADLAKYOKDL-ADYPVKLKAYEDEQASIKAAALELKHKNED--GNLTPTS 504
QY	516 LTRFSVPRATEARGPEKTDQALARVISGNSRTGLDVSQVSAL-----QSVMQ-----563
Db	505 AQNLVYDLEPNANUSLTDDGKFLK-----ASAVDDAFSKSTSKAKYDQKILQLDDLDITN 559
QY	564 ITCQSPQANNEEI-----RQKLTSAVTKPPQFGYPYVQLS-NDSTQKFTAKLESFAEG 616
Db	560 LEQSDNVASSMELYNFGCDKAGNETTYSNNSQVKKWSVLLERQGSATATYTNLQNSVYNG 619
QY	617 SRTAAEIKALSFETNS-----LFIQOVLVNIIGLSYGYLK 651
Db	620 KKISKIVKYTYDPKSKFOGOKWLGI:FTDPTLGVFASAYTGVE 664

```

RESULT 12
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
; APPLICANT: Genecor
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394-PCT
; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRF
; ORGANISM: Bacillus subtilis
US-09-308-375-2

Query Match 5.4%; Score 173; DB 4; Length 2285;
Best Local Similarity 20.4%; Pred. No. 0.003;
Matches 171; Conservative 126; Mismatches 292; Indels 250; Gaps 37;

QY 20 DLSAOGLE--ASAANKSAEQAORIA---GAEAKPKESKTD---SVERWSILRGAVNALMSL 71
DB DLDPELEKFSLSGLKLOEMKALDSGDEKAFDAKQLOSLLETYSKSDSIDVFKMS 1040
QY 72 ADKL--GIASSNSSSTSR--ADVDSTTATPTPPPTFDYKQATPAYDTITSTSL 127
DB 1041 FDKAQKNIKDGKSLSSVKSEVSGDLGTLAEAGNEA---EDFGKKEALD---ANSV 1092
QY 128 ADIQAALVSLQAV---TNIKD-----TAATDEETAI 156
DB 1093 DDIAKAIKEMSDAMQFSDVQVLNGDFNNTKQVAPLNDLLEKMAEGKSIANEANTLI 1152
QY 157 AAEMETKNADAVKGA---QITELAKYASD-----NOAI----- 187
DB 1153 QKDKELAQALSIENGVVVKNRDEVIKQKVKLDAYNDVVTYSNKLMTKTEVNAIKTLNAD 1212
QY 188 ---LDSGLK---TSPDLQAL---LOSANNKKAELLKEMODNPVVP 229
DB 1213 TLRLDLSLKKLKERKLDMSAEISDLEVKSIINNADAKKELKLEKMLQPGGYSNQIE 1272
QY 230 ---KTPALAQSLVDOTATATQIEKDGNAIRDAYFAGQNASGAVE---NAKSNNSISNTD 283
DB 1273 AMQSVKSALESYISABEATSTQ--EMNKQALVEAGTSLNWTDDQKABEETKTSMYVD 1331
QY 284 SAKAALATATQIAEAKPKFPDPS-----ILOAEQMWIOAEKDLKN--- 325
DB 1332 KYKEALEKYNAEIDKYNQVNDYKPYKSYKQYDAIKKEIKALQOKKMLQEAQKLLKQK 1391
QY 326 -----TKPADGSDVPNGPTVVG--SKQGGSSIGSIRVSMILLDAENETASILMSGFR 376
DB 1392 SGNITQYGIWTTSTSSGGTSPSTGGSYSGKYSYINSAASKYNVDPA-----LIAAVI 1444
QY 377 OMIHMENTENPDSQAQOELAAQARAARAKAGDDSAALADAO---KALEAALGK--AGQ 432
DB 1445 QOESGENAKARSGVAGMGLQMLPEATAKSLGVNNAYDPYONVWVGKYLQAQLEKFGNV 1504
QY 433 QGILNALGOIASAAVVSAGVPP-----AAASIGSVKQLY----- 468
DB 1505 EKALAAYNAGPGNVIKYGGIPPPKETQNYVKIMANYKSLSATSSIASYYTNSAFRV 1564
QY 469 -----KTSKSTGSDY-----KTQISAGVDAYKSIDAYGRARND 503
DB 1565 SSKYGOQESGLRSPHKGTDFAAKAGTAKTSLQSGKVQI--AGYS--KTAGNWWVIKQDDG 1621
QY 504 T-RDVNNVSTPALT--RSVPRART-----EARG-----PEKTDQAL 537
DB 1622 TVAKYHMLNTPSKAGQSVKAGQTTGKVGSTGNSTGNHHLQIENGKTIIDPKYMQGI 1681
QY 538 ARVISGNSRTGLGVYQVSALQVMOIIOINPOANNEIRQKILTSVATKPPQFGYPVQL 597
DB 1682 GTSISDASQAERQOQIAQAKSDLLSLQGDISSVNDIQ-----ELQYELVQS 1730
QY 598 SNDSTOK-----FIAKLSL---FAEGR-----TAARIKALSRETSLFTQQVLVN 641
DB 1731 KLDEFDKRIGDFDVRIAKDESMANRYTSDSKFRKVTSDQKAVABQAK---IQOQKVN 1786

RESULT 13

US-09-540-236-2753
; Sequence 2753, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

RESULT 14

US-08-687-956A-23
; Sequence 23, Application US/08687956A
; Patent No. 5861157

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAF
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2753
; LENGTH: 718
; TYPE: PRF
; ORGANISM: M. catarrhalis
US-09-540-236-2753

Query Match 5.3%; Score 170.5; DB 4; Length 718;

Best Local Similarity 19.2%; Pred. No. 0.0089;

Matches 144; Conservative 120; Mismatches 248; Indels 237; Gaps 36;

QY 11 DETERTPPADLSAQL-----EASRANKSAEQR-----IAGAEAKPKESKTD 53
DB 32 DDVEIVAEVDIDNQAIVSDVLIIRDITDKADQADHTDDASKADDETVDGVKQKQAEKED 91
QY 54 SVERWSILRSVAVNALMSLADKLGIIASSNSSSTSRSDVDSTATATPTPPPTFDYKQ 113
DB 92 FENKAQDLQD-----KATEKLEVAKEATQDKVEKTSQSL-----VEDIKOK 131
QY 114 AQATYDTTFTSLADIQAALVSLQDAYTNIKDTAATDEETALAFAEWETKNADAVKGAQ 173
DB 132 AQSLQE-----DAADTVEALKQ--AASDKVETTKAEQSLKDDA---TQ 170
QY 174 TTELAKYASD-----NOAILDLSGLK-----TSPD-----LLOAALQSVANNKA 214
DB 171 TFSAKQAVEGKVEAIKQVLDQVDSLKDDTDODNTDQDQEQKTLKDXAVQAATAAKRV 230
QY 215 AELLKEMODNP-----VVPKTPALIAQSLVDOTATATQIEKDGNAIRDAYFAGQNASGA 269
DB 231 EDVDDVXHTTSPKNTASGKIDKQAAVDKTEBVKSQLSQKADALKS---SGBELKQT 287
QY 270 VENAKSNNSISNIDSAKAIATAKTQIAEAKPKFPDPSILOAEQMWIOAEKDLKNI--- 326
DB 288 ACTA--ANDAIT---EAQAAVVGSGVAAA-----DSAQSTQAQAKDKLNQLE 330
QY 327 --KPADGSDVPNGPTTVGSGKQGGSSIGSIRVSMILLDAENETASILMSGFRQMIHMENT 384
DB 331 QGKSALDEKVOELGEKFGATE-----KINAVSENVDLATQVI----- 367
QY 385 ENPDSQAQOELAAQARAARAKAGDDSAALADAOALEAALGKAG----- 430
DB 368 -KEEAQALQTNAAQESLQAAKAAGEBYDA---THEDKGLTTLKLGKVGAVLSGMYGISONKN 423
QY 431 -QOQGI-----LNA-----LGOIASAAVVSAG-----VPPAAASIGSSVKOL 467
DB 424 KHYQGVDLHRSFDDKDAFHAQSSFFAGQIFGAKAVAKNVAKVFPQSFKEAIGES---L 480
QY 468 YKTSKSTGSDYKTOISAGYDAYKSIDAYGRARNDATRDVINNVSTPALTRSVPRARTEA 527
DB 481 Y-----NKVAESNMA--AIKD---LKNPRFDLINTMT-----QERHAPA 517
QY 528 RGEPTKDOALRV-----ISGNSRTGLD--VYSQVSALQSVMOI---IQSNPOANNEIROK 579
DB 518 EDVANQNRALATLGGVAGLAGLKGVLADAAMLLMVLSTRTVYQVAAIYDQPLTGKEGKTK- 576
QY 580 LTSVATKPPQFGYPVQVLSNDSTQKF-----LAKLESIFAEGSRTA--AEIKALS- 627
DB 577 -----AYGVLSGANLEKQKQVILTALGSSMLANAQQTGIKAQLDLSLA 623
QY 628 -FETNSLFTQQVL-----VNIQSLYSGYL 650
DB 624 RYRESQPIYAKQFLLDKFVNLDNLNPNML 652

```

; GENERAL INFORMATION:
; APPLICANT: BURNIE, JAMES P
; APPLICANT: MATTHEWS, RUTH C
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: INFECTIONS DUE TO STREPTOCOCCI AND ENTEROCOCCI
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, LLP
; STREET: 1100 NEW YORK AVENUE, N.W., EAST TOWER, 9TH
; STREET: FLOOR
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,956A
; FILING DATE: 29-JUL-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9401689.6
; FILING DATE: 28-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 50885/222892
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/861-3000
; TELEFAX: 202/822-0944
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus sobrinus
; US-08-687-956A-23

```

```

Query Match 5.2%; Score 167; DB 2; Length 1566;
Best Local Similarity 21.7%; Pred. No. 0.0046;
Matches 131; Conservative 91; Mismatches 245; Indels 136; Gaps 22;

Qy 48 KESKTSVERWSILRSVAVNLSLADKGLIASSNSSSTS--RSADVSTTATPTPTPP 105
Db 10 RKSISRITLGLGTAI--LASVTGKALAEETSTTSGVNTAVVGTETGNPATNLPD 67
Qy 106 TFDYKTAQATY-----DTFTSLADIQAALVSLQDAVINIKDTATD----- 151
Db 68 KQNPSSQAETSQAQAGCKTGAMSDVSTSELDEAKSAQAGVTVSQDATVDKGVETS 127
Qy 152 -----EETATAAEWETKNADAVKGAQITELAKYA--SDNQAILDSLGLTSFLLQALL 205
Db 128 DEANQKETEIKDDYSKQADIQKT-----TEYKKAVERNQAETD---RITQENAKKAQY 180
Qy 206 QSVANNKAELLKEMQNPVVPKTPAIAQSLVDQTDATATQIEKQGNIRDAYFAGQN 265
Db 181 EQDLAANK-AEVERTITNEN-----AQKADY-EAKLAQYQKDLAAVQQAANDSQA 228
Qy 266 ASGAVENAKSNNSISNIDSAKAATATAKTQIAEA-----QKKEPDSILO----- 310
Db 229 AYAAAKEA-YDKELARVQAANAA--AKKEYEALANTTNEQIKAEANAIQQRNAQK 284
Qy 311 -EAEQWITQAEKDLKLNKIPADGSDVPNGTIVGSGKQSGSIGSVIRSMILLDDAENETAS 369

```

```

Db 285 ADYEAKLAQYKDL-----AAAQSG-----NATNE----- 309
Qy 370 ILMGFRQMIHMFNTENPDQAQOELAAQARAAXAGDDSAALAA--DAQKALEAALG 427
Db 310 -----ADYQAKKAAYQELARVQAANAAKQAVEQALAAANTAKNAQITAE 355
Qy 428 KAGQQGIIINALGQITASAAVWSAGVPPAAASIGSSVKOLYKTSKSTGSDYKTQISAGYD 487
Db 356 EAIQQR--NAQAKANYEAKLAQYQKDLAAQSGNAANE-----ADYQEKLAAYEK 403
Qy 488 AVKSNIDAYGRANDATRDVINNVSTPALTRSPARTAEARGPEKTDQALARVIGNSRT 547
Db 404 ELARVQAANAAKQYEQVQVEANAGNAEITANRAIRERNNAKAKTDYELK--LSKYQEE 461
Qy 548 LGDVYSQVSALOSVMQIIQSNFPANNEIRQKLTSAVTKPPQFGYPYVOLSDNDSTQKFTA 607
Db 462 LAQYKDLAEYPAKLQAYQEQAA-----IKAALELEKHKNEEDWNLSPSAQSLVY 513
Qy 608 KLE 610
Db 514 DLE 516

RESULT 15
US-09-874-923-120
; Sequence 120, Application US/09874923
; Patent No. 6638517
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhatia, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; FILE REFERENCE: 210121.420C8
; CURRENT APPLICATION NUMBER: US/09/874,923
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 2310
; TYPE: PRT
; ORGANISM: Leishmania major and chagasi
US-09-874-923-120

```

```

Query Match 5.2%; Score 165.5; DB 4; Length 2310;
Best Local Similarity 19.9%; Pred. No. 0.01;
Matches 134; Conservative 122; Mismatches 295; Indels 121; Gaps 24;

Qy 12 ETERTPPADLSAAGLEASAANKSAEQAIRIAGAEKESKTSVERWSILRSVAVNLSL 71
Db 984 DTATQORAELEAQLARLAADREARQQLAANAEELQORLDTATQOR-AELEAQLAR 1042
Qy 72 ADKLIASSNSSSTSRSADVSTTATPTPTPPPTFDYKTAQATAYDTFTST-SLADI 130
Db 1043 ABEEL---QORLDTATQORAELEARVARLAADREARQQLAANAEELQORLDTATQOR 1099
Qy 131 QAAALVSLQDAVINIKDTATDEETATAAEWEYKNADAVKGAQITELAKYASNDQAILDS 190
Db 1100 EARVARL-----AANAEE--LQORLDTATQORAELEAQLARLAANAEELQORLDT 1147
Qy 191 LGLKTSFLLQALLQSVANNKKAELL---KEMQNPVVPKTPAIAQSLVDQTDATA 246
Db 1148 ATQORA--ELEARVARLAADREARQQLAANAEELQORLDTATQORAELEAQLARLAANA 1205
Qy 247 TQTEK-DGNAIRDAYFAGQNAAGAVENAKSNNSI--SNIDSAKAATATAKTQIAEAQK 303
Db 1206 EELQORLDTATQORAELEAQLARLAADREARQQLAANAEELQORLDTATQORAELE--- 1262

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 05:59:21 ; Search time 152.132 Seconds
(without alignments)
1108.119 Million cell updates/sec

Title: US-10-608-559-2
Perfect score: 3187
Sequence: 1 MVNPIGPGPIDETERTPPAD.....SLFIQVLVNIISLYSGYLQ 651

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	9	US-09-841-132-405
2	3187	100.0	651	10	US-09-564-479-2
3	3187	100.0	651	10	US-09-841-260-78
4	3187	100.0	651	13	US-10-007-693-78
5	3187	100.0	651	15	US-10-312-273-13
6	2844	89.2	583	10	US-09-564-479-4
7	2208	69.3	490	10	US-09-564-479-8
8	2204	69.2	452	10	US-09-564-479-6
9	2182	68.5	478	15	US-10-289-762-776
10	1000	31.4	212	15	US-10-289-762-775
11	899.5	28.2	647	9	US-09-841-132-436
12	894	28.1	660	10	US-09-841-260-139
13	894	28.1	660	13	US-10-007-693-139
14	671.5	21.1	361	9	US-09-841-132-299
15	251	7.9	1463	9	US-09-971-536-69

16	235.5	7.4	1178	14	US-10-128-714-9240	Sequence 8240, Ap
17	234	7.3	1831	12	US-10-282-122A-71033	Sequence 71033, A
18	231.5	7.3	2478	9	US-09-815-242-5816	Sequence 5816, Ap
19	231.5	7.3	2478	9	US-09-815-242-12967	Sequence 12967, A
20	223.5	7.0	2481	12	US-10-282-122A-43762	Sequence 43762, A
21	218	6.8	603	12	US-10-282-122A-72004	Sequence 72004, A
22	217.5	6.8	6713	12	US-10-282-122A-43811	Sequence 43811, A
23	217	6.8	5795	9	US-09-815-242-12310	Sequence 12610, A
24	213	6.7	2434	9	US-09-815-242-5835	Sequence 5835, Ap
25	213	6.7	2434	9	US-09-815-242-12996	Sequence 12996, A
26	210	6.6	6122	9	US-10-282-122A-42783	Sequence 42783, A
27	210	6.6	1123	12	US-10-282-122A-70581	Sequence 70581, A
28	210	6.6	2271	12	US-10-282-122A-43924	Sequence 43924, A
29	208.5	6.5	4776	12	US-10-282-122A-73678	Sequence 73678, A
30	207.5	6.5	2344	9	US-09-815-242-12713	Sequence 12713, A
31	207	6.5	3533	12	US-10-282-122A-70177	Sequence 70177, A
32	204	6.4	2283	14	US-10-172-502-4	Sequence 4, Appl1
33	203.5	6.4	2368	9	US-09-815-242-5635	Sequence 5635, Ap
34	203.5	6.4	2368	9	US-09-815-242-12389	Sequence 12389, A
35	201.5	6.3	1879	9	US-09-971-536-70	Sequence 70, Appl1
36	201	6.3	459	12	US-10-282-122A-70453	Sequence 70453, A
37	201	6.3	1142	15	US-10-369-493-10778	Sequence 10778, A
38	199.5	6.3	2025	9	US-09-815-242-5703	Sequence 5703, Ap
39	199.5	6.3	3158	9	US-09-815-242-12611	Sequence 12611, A
40	199	6.2	6641	12	US-10-282-122A-70580	Sequence 70580, A
41	198.5	6.2	1215	9	US-09-815-242-5908	Sequence 5908, Ap
42	198.5	6.2	1269	9	US-09-815-242-13113	Sequence 13113, A
43	198	6.2	1222	12	US-10-282-122A-69799	Sequence 69799, A
44	196	6.1	560	9	US-09-815-242-13057	Sequence 13057, A
45	196	6.1	1924	12	US-10-282-122A-71846	Sequence 71846, A

ALIGNMENTS

RESULT 1
US-09-841-132-405
; Sequence 405, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 405
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-405

Query Match	100.0%	Score	3187	DB	9	Length	651
Best Local Similarity	100.0%	Pred. No.	4e-225				
Matches	651	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MVNPIGPGPIDETERTPPADLSAQLSAAKSAQAQIAAGAKPKESKTSVERMSI	60				
Db	1	MVNPIGPGPIDETERTPPADLSAQLSAAKSAQAQIAAGAKPKESKTSVERMSI	60				
Qy	61	LRSAVNALMSLADKLGIIASSNSSSTSRSDVDSVDTTATPTPPPTDDYKTAQAYDT	120				
Db	61	LRSAVNALMSLADKLGIIASSNSSSTSRSDVDSVDTTATPTPPPTDDYKTAQAYDT	120				
Qy	121	IFTSTSLADIQAALVSLQDAVTNKTDAATDEETAAAEWETKNADAVKVCQAOTELAKY	180				
Db	121	IFTSTSLADIQAALVSLQDAVTNKTDAATDEETAAAEWETKNADAVKVCQAOTELAKY	180				
Qy	181	ASDNCAILDLSGLKTSFDLLQAALLQS VANNKAAELLKEMQDNFVPGKTPAIAQSLVD	240				

Db 181 ASDNQAILDSGLKTSFDLLOAALQSVAANNKAAELLKEMQDNVPVPGKTPALQAQSLVD 240
Qy 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAALATAKTQIAEAQ 300
Db 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAALATAKTQIAEAQ 300
Qy 301 KKFPPSPILQEAQEMVIOAEKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSRVSMML 360
Db 301 KKFPPSPILQEAQEMVIOAEKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSRVSMML 360
Qy 361 DDAENETASILMSGFRQMIHNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
Db 361 DDAENETASILMSGFRQMIHNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
Qy 421 ALEAALGKAGQOQOGLNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKT 480
Db 421 ALEAALGKAGQOQOGLNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKT 480
Qy 481 QISAGYDAYKSIDAYGRANDATRDVINNVSTPALTRSPRARTTEARGPEKTDQALARV 540
Db 481 QISAGYDAYKSIDAYGRANDATRDVINNVSTPALTRSPRARTTEARGPEKTDQALARV 540
Qy 541 ISGNSRTLGDVYSQVSALOSVMQIIQSNPQANNEEIRQKLTSAVTKPPQFGYPYVQLSND 600
Db 541 ISGNSRTLGDVYSQVSALOSVMQIIQSNPQANNEEIRQKLTSAVTKPPQFGYPYVQLSND 600
Qy 601 STQKFIKLESIFAEGSRTAABEIKALSFEFNSLFIQOVLVNGSLYSGLYQ 651
Db 601 STQKFIKLESIFAEGSRTAABEIKALSFEFNSLFIQOVLVNGSLYSGLYQ 651

RESULT 2
US-09-564-479-2
; Sequence 2, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COHEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 651
; TYPE: PRF
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-2

Query Match 100.0%; Score 3187; DB 10; Length 651;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSABEORAGAEAKPKESKTDVSVERWSI 60
Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSABEORAGAEAKPKESKTDVSVERWSI 60
Qy 61 LRSVAVNALSADKLGIASSNSSSTSRSDVSDTATATPTPPPTFDYKTOQATAYDT 120
Db 61 LRSVAVNALSADKLGIASSNSSSTSRSDVSDTATATPTPPPTFDYKTOQATAYDT 120
Qy 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180

Qy 181 ASDNQAILDSGLKTSFDLLOAALQSVAANNKAAELLKEMQDNVPVPGKTPALQAQSLVD 240
Db 181 ASDNQAILDSGLKTSFDLLOAALQSVAANNKAAELLKEMQDNVPVPGKTPALQAQSLVD 240
Qy 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAALATAKTQIAEAQ 300
Db 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAALATAKTQIAEAQ 300
Qy 301 KKFPPSPILQEAQEMVIOAEKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSRVSMML 360
Db 301 KKFPPSPILQEAQEMVIOAEKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSRVSMML 360
Qy 361 DDAENETASILMSGFRQMIHNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
Db 361 DDAENETASILMSGFRQMIHNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
Qy 421 ALEAALGKAGQOQOGLNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKT 480
Db 421 ALEAALGKAGQOQOGLNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKT 480
Qy 481 QISAGYDAYKSIDAYGRANDATRDVINNVSTPALTRSPRARTTEARGPEKTDQALARV 540
Db 481 QISAGYDAYKSIDAYGRANDATRDVINNVSTPALTRSPRARTTEARGPEKTDQALARV 540
Qy 541 ISGNSRTLGDVYSQVSALOSVMQIIQSNPQANNEEIRQKLTSAVTKPPQFGYPYVQLSND 600
Db 541 ISGNSRTLGDVYSQVSALOSVMQIIQSNPQANNEEIRQKLTSAVTKPPQFGYPYVQLSND 600
Qy 601 STQKFIKLESIFAEGSRTAABEIKALSFEFNSLFIQOVLVNGSLYSGLYQ 651
Db 601 STQKFIKLESIFAEGSRTAABEIKALSFEFNSLFIQOVLVNGSLYSGLYQ 651

RESULT 3
US-09-841-260-78
; Sequence 78, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121-515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 78
; LENGTH: 651
; TYPE: PRF
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-78

Query Match 100.0%; Score 3187; DB 10; Length 651;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSABEORAGAEAKPKESKTDVSVERWSI 60
Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSABEORAGAEAKPKESKTDVSVERWSI 60
Qy 61 LRSVAVNALSADKLGIASSNSSSTSRSDVSDTATATPTPPPTFDYKTOQATAYDT 120
Db 61 LRSVAVNALSADKLGIASSNSSSTSRSDVSDTATATPTPPPTFDYKTOQATAYDT 120
Qy 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Qy 181 ASDNQAILDSGLKTSFDLLOAALQSVAANNKAAELLKEMQDNVPVPGKTPALQAQSLVD 240
Db 181 ASDNQAILDSGLKTSFDLLOAALQSVAANNKAAELLKEMQDNVPVPGKTPALQAQSLVD 240

QY 241 QTDAATQIEKDGNAIRDAYFAGONASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQ 300
 DB 241 QTDAATQIEKDGNAIRDAYFAGONASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQ 300
 QY 301 KFPDPSPILOEABQWIOAEKDLKNIKPADGSDVNPFGTTVGGSKOQSSIGSIRVSMML 360
 DB 301 KFPDPSPILOEABQWIOAEKDLKNIKPADGSDVNPFGTTVGGSKOQSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMENTENPDSSQAQOELAAQARAARAAAGDDSSAAAALADAQK 420
 DB 361 DDAENETASILMSGFRQMIHMENTENPDSSQAQOELAAQARAARAAAGDDSSAAAALADAQK 420
 QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAAASIGSSVKQLYKTSKSTGSDYKT 480
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAAASIGSSVKQLYKTSKSTGSDYKT 480
 QY 481 QISAGYDAYKSINDAYGRANDATROVINNVSTPALTRSPRARTARGPEKTDQALAV 540
 DB 481 QISAGYDAYKSINDAYGRANDATROVINNVSTPALTRSPRARTARGPEKTDQALAV 540
 QY 541 ISGNSRTLGDVYSQVSALQSVMOIIOQSNPOANNEEIRQKLTSAVTKPPQFGYPPVQLSND 600
 DB 541 ISGNSRTLGDVYSQVSALQSVMOIIOQSNPOANNEEIRQKLTSAVTKPPQFGYPPVQLSND 600
 QY 601 STQKFIKLESLEFAGSRTAAEIKALSFTETNSLFIQOVLVNIIGSLYSGYLQ 651
 DB 601 STQKFIKLESLEFAGSRTAAEIKALSFTETNSLFIQOVLVNIIGSLYSGYLQ 651
 RESULT 4
 US-10-007-693-78
 ; Sequence 78, Application US/10007693
 ; Publication No. US20020146776A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
 ; FILE REFERENCE: 210121.515C2
 ; CURRENT APPLICATION NUMBER: US/10/007,693
 ; CURRENT FILING DATE: 2001-12-05
 ; NUMBER OF SEQ ID NOS: 157
 ; SEQ ID NO 78
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-10-007-693-78

Query Match 100.0%; Score 3187; DB 13; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4e-225;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNPPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
 DB 1 MNPPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
 QY 61 LRSVAVNALMSLADKLGIASSNSSSSSTSRSDVDSTTATPTPPPTDDYKTAQATYDT 120
 DB 61 LRSVAVNALMSLADKLGIASSNSSSSSTSRSDVDSTTATPTPPPTDDYKTAQATYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVTNICKDTAADEETAIAAEWETKNADAVKGAQITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTNICKDTAADEETAIAAEWETKNADAVKGAQITELAKY 180
 QY 181 ASDNQATILDSIGKLTSDLLQAAQLQSVANNKKAELLKEMQDNPVPGKTPPAQAQLSVD 240
 DB 181 ASDNQATILDSIGKLTSDLLQAAQLQSVANNKKAELLKEMQDNPVPGKTPPAQAQLSVD 240
 QY 241 QTDAATQIEKDGNAIRDAYFAGONASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQ 300
 DB 241 QTDAATQIEKDGNAIRDAYFAGONASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQ 300

QY 301 KFPDPSPILOEABQWIOAEKDLKNIKPADGSDVNPFGTTVGGSKOQSSIGSIRVSMML 360
 DB 301 KFPDPSPILOEABQWIOAEKDLKNIKPADGSDVNPFGTTVGGSKOQSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMENTENPDSSQAQOELAAQARAARAAAGDDSSAAAALADAQK 420
 DB 361 DDAENETASILMSGFRQMIHMENTENPDSSQAQOELAAQARAARAAAGDDSSAAAALADAQK 420
 QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAAASIGSSVKQLYKTSKSTGSDYKT 480
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAAASIGSSVKQLYKTSKSTGSDYKT 480
 QY 481 QISAGYDAYKSINDAYGRANDATROVINNVSTPALTRSPRARTARGPEKTDQALAV 540
 DB 481 QISAGYDAYKSINDAYGRANDATROVINNVSTPALTRSPRARTARGPEKTDQALAV 540
 QY 541 ISGNSRTLGDVYSQVSALQSVMOIIOQSNPOANNEEIRQKLTSAVTKPPQFGYPPVQLSND 600
 DB 541 ISGNSRTLGDVYSQVSALQSVMOIIOQSNPOANNEEIRQKLTSAVTKPPQFGYPPVQLSND 600
 QY 601 STQKFIKLESLEFAGSRTAAEIKALSFTETNSLFIQOVLVNIIGSLYSGYLQ 651
 DB 601 STQKFIKLESLEFAGSRTAAEIKALSFTETNSLFIQOVLVNIIGSLYSGYLQ 651
 RESULT 5
 US-10-312-273-13
 ; Sequence 13, Application US/10312273
 ; Publication No. US20040005667A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SPA
 ; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
 ; FILE REFERENCE: P02503SWO
 ; CURRENT APPLICATION NUMBER: US/10/312,273
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: 0016363.4
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 0017047.2
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 0017983.8
 ; PRIOR FILING DATE: 2000-07-21
 ; PRIOR APPLICATION NUMBER: 0019368.0
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: 0020440.4
 ; PRIOR FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: 0022583.9
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 0027549.5
 ; PRIOR FILING DATE: 2000-11-10
 ; PRIOR APPLICATION NUMBER: 0031706.5
 ; PRIOR FILING DATE: 2000-12-22
 ; NUMBER OF SEQ ID NOS: 664
 ; SOFTWARE: SeqWin99, version 1.02
 ; SEQ ID NO 13
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-10-312-273-13

Query Match 100.0%; Score 3187; DB 15; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4e-225;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNPPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
 DB 1 MNPPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
 QY 61 LRSVAVNALMSLADKLGIASSNSSSSSTSRSDVDSTTATPTPPPTDDYKTAQATYDT 120
 DB 61 LRSVAVNALMSLADKLGIASSNSSSSSTSRSDVDSTTATPTPPPTDDYKTAQATYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVTNICKDTAADEETAIAAEWETKNADAVKGAQITELAKY 180

Db 121 IFTSTSLADIQAALVSLQDAVTNINKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 Qy 181 ASDNQAILDSLGKTSFLLQALLOSANNKAAELLKEMQDNPVPGKTPALQAQSLVD 240
 Db 181 ASDNQAILDSLGKTSFLLQALLOSANNKAAELLKEMQDNPVPGKTPALQAQSLVD 240
 Qy 241 QTDATATQIEKDGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
 Db 241 QTDATATQIEKDGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
 Qy 301 KKFPPSPILQBAQVQVIAEKDLKNIKPADGSDVENPGTTVGGSKQOQSSISIRVSMIL 360
 Db 301 KKFPPSPILQBAQVQVIAEKDLKNIKPADGSDVENPGTTVGGSKQOQSSISIRVSMIL 360
 Qy 361 DDAENETASILMSGFRMHMENTENPDPSQAAQELAAQAAKAAAGDSDSAAALADAQK 420
 Db 361 DDAENETASILMSGFRMHMENTENPDPSQAAQELAAQAAKAAAGDSDSAAALADAQK 420
 Qy 421 ALEAALGKAGQOQGIINLALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 Db 421 ALEAALGKAGQOQGIINLALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 Qy 481 QISAGDYKSNIDAYGRANDATRDVINNVSTPALTRSVPRARTEARPEKTDQALARV 540
 Db 481 QISAGDYKSNIDAYGRANDATRDVINNVSTPALTRSVPRARTEARPEKTDQALARV 540
 Qy 541 ISGNSRTLGDVYSQVSAQVSMQIITQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
 Db 541 ISGNSRTLGDVYSQVSAQVSMQIITQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
 Qy 601 STQKFIKLESFAEGSRTAAEIKALSFETNSLFIQQVLVNTGSLYSGYLQ 651
 Db 601 STQKFIKLESFAEGSRTAAEIKALSFETNSLFIQQVLVNTGSLYSGYLQ 651

RESULT 6

US-09-564-479-4

; Sequence 4, Application US/09564479

; Publication No. US2003095973A1

; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: COMEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND

; FILE REFERENCE: US/09/564,479

; CURRENT APPLICATION NUMBER: US/09/564,479

; PRIOR FILING DATE: 2000-05-03

; PRIOR FILING DATE: 1999-05-03

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 583

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-564-479-4

Query Match

Best Local Similarity 99.8%; Score 2844; DB 10; Length 583;

Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 69 MSADKLGIASSSSSTSSADVSDSTATAPTPPPPTFDYKTAQYDTITFTSLA 128
 Db 1 MSADKLGIASSSSSTSSADVSDSTATAPTPPPPTFDYKTAQYDTITFTSLA 60
 Qy 129 DIQAALVSLQDAVTNINKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKYSDNQAIL 188
 Db 61 DIQAALVSLQDAVTNINKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKYSDNQAIL 120

Qy 189 DSLGKTSFLLQALLOSANNKAAELLKEMQDNPVPGKTPALQAQSLVDQTDATQ 248
 Db 121 DSLGKTSFLLQALLOSANNKAAELLKEMQDNPVPGKTPALQAQSLVDQTDATQ 180
 Qy 249 IEKDGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQKFFPDSPI 308
 Db 181 IEKDGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQKFFPDSPI 240
 Qy 309 LQEAQVQVIAEKDLKNIKPADGSDVENPGTTVGGSKQOQSSISIRVSMILDDAENETA 368
 Db 241 LQEAQVQVIAEKDLKNIKPADGSDVENPGTTVGGSKQOQSSISIRVSMILDDAENETA 300
 Qy 369 SILMSGFRMHMENTENPDPSQAAQELAAQAAKAAAGDSDSAAALADAQKALEAALGK 428
 Db 301 SILMSGFRMHMENTENPDPSQAAQELAAQAAKAAAGDSDSAAALADAQKALEAALGK 360
 Qy 429 AGQOQGIINLALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKTQISAGYDA 488
 Db 361 AGQOQGIINLALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKTQISAGYDA 420
 Qy 489 YKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARPEKTDQALARVISGNSRTL 548
 Db 421 YKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARPEKTDQALARVISGNSRTL 480
 Qy 549 GDVYSQVSAQVSMQIITQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 608
 Db 481 GDVYSQVSAQVSMQIITQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 540
 Qy 609 LESLFAEGSRTAAEIKALSFETNSLFIQQVLVNTGSLYSGYLQ 651
 Db 541 LESLFAEGSRTAAEIKALSFETNSLFIQQVLVNTGSLYSGYLQ 583

RESULT 7

US-09-564-479-8

; Sequence 8, Application US/09564479

; Publication No. US20030095973A1

; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: COMEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND

; FILE REFERENCE: US/09/564,479

; CURRENT APPLICATION NUMBER: US/09/564,479

; PRIOR FILING DATE: 2000-05-03

; PRIOR FILING DATE: 1999-05-03

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 490

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-564-479-8

Query Match

Best Local Similarity 99.8%; Score 2208; DB 10; Length 490;

Matches 453; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVNPIGPGPIDETETPPADLSAQGLEASAAKSAEORIAAGAEAKPKESKTDSEVERSI 60
 Db 1 MVNPIGPGPIDETETPPADLSAQGLEASAAKSAEORIAAGAEAKPKESKTDSEVERSI 60
 Qy 61 LRSVNNALMSLADKLGIASSSSSTSSADVSDSTATAPTPPPPTFDYKTAQYDT 120
 Db 61 LRSVNNALMSLADKLGIASSSSSTSSADVSDSTATAPTPPPPTFDYKTAQYDT 120
 Qy 121 IFTSTSLADIQAALVSLQDAVTNINKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180

Db 121 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVAQITELAKY 180
 Qy 181 ASDNQAALDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNPVVPKTPAIAQSLVD 240
 Db 181 ASDNQAALDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNPVVPKTPAIAQSLVD 240
 Qy 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSINIDSAAKAIATAKTQIAEAQ 300
 Db 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSINIDSAAKAIATAKTQIAEAQ 300
 Qy 301 KKFPDPSPIQAEQWVIQAEKDLKNIKPAGSDVPNPGTTVGGSKQGGSSIGSVSMML 360
 Db 301 KKFPDPSPIQAEQWVIQAEKDLKNIKPAGSDVPNPGTTVGGSKQGGSSIGSVSMML 360
 Qy 361 DDAENETASILMSGFQMIHMFNTENPDSSQAQOELAAQAAKAGDSSAAALADAQK 420
 Db 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 454
 Db 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 454

RESULT 8

US-09-564-479-6
 ; Sequence 6, Application US/09564479
 ; Publication No. US2003009573A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: OOMEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; CURRENT FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/132,270
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR APPLICATION NUMBER: 60/141,276
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 452
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-564-479-6

Query Match 69.2%; Score 2204; DB 10; Length 452;
 Best Local Similarity 100.0%; Pred. No. 3.le-153;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTDVSVERWSI 60
 Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTDVSVERWSI 60
 Qy 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSDTTATPTPPPTTDDYKTAQTAAYDT 120
 Db 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSDTTATPTPPPTTDDYKTAQTAAYDT 120
 Qy 121 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVAQITELAKY 180
 Db 121 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVAQITELAKY 180
 Qy 181 ASDNQAALDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNPVVPKTPAIAQSLVD 240
 Db 181 ASDNQAALDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNPVVPKTPAIAQSLVD 240
 Qy 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSINIDSAAKAIATAKTQIAEAQ 300
 Db 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSINIDSAAKAIATAKTQIAEAQ 300

Qy 301 KKFPDPSPIQAEQWVIQAEKDLKNIKPAGSDVPNPGTTVGGSKQGGSSIGSVSMML 360
 Db 301 KKFPDPSPIQAEQWVIQAEKDLKNIKPAGSDVPNPGTTVGGSKQGGSSIGSVSMML 360
 Qy 361 DDAENETASILMSGFQMIHMFNTENPDSSQAQOELAAQAAKAGDSSAAALADAQK 420
 Db 361 DDAENETASILMSGFQMIHMFNTENPDSSQAQOELAAQAAKAGDSSAAALADAQK 420
 Qy 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452
 Db 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452

RESULT 9

US-10-289-762-776
 ; Sequence 776, Application US/10289762
 ; Publication No. US20040006218A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffiths, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/10/289,762
 ; CURRENT FILING DATE: 2003-03-27
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 776
 ; LENGTH: 478
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-10-289-762-776

Query Match 68.5%; Score 2182; DB 15; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.4e-151;
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTDVSVERWSI 60
 Db 3 MWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTDVSVERWSI 62
 Qy 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSDTTATPTPPPTTDDYKTAQTAAYDT 120
 Db 63 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSDTTATPTPPPTTDDYKTAQTAAYDT 122
 Qy 121 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVAQITELAKY 180
 Db 123 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVAQITELAKY 182
 Qy 181 ASDNQAALDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNPVVPKTPAIAQSLVD 240
 Db 183 ASDNQAALDSGLKTSFDLLQAALQSVANNKAAELLKEMQDNPVVPKTPAIAQSLVD 242
 Qy 241 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSINIDSAAKAIATAKTQIAEAQ 300
 Db 243 QTDATATQIEKGNNAIRDYFAGQNASGAVENAKSNNSINIDSAAKAIATAKTQIAEAQ 302
 Qy 301 KKFPDPSPIQAEQWVIQAEKDLKNIKPAGSDVPNPGTTVGGSKQGGSSIGSVSMML 360
 Db 303 KKFPDPSPIQAEQWVIQAEKDLKNIKPAGSDVPNPGTTVGGSKQGGSSIGSVSMML 362
 Qy 361 DDAENETASILMSGFQMIHMFNTENPDSSQAQOELAAQAAKAGDSSAAALADAQK 420
 Db 363 DDAENETASILMSGFQMIHMFNTENPDSSQAQOELAAQAAKAGDSSAAALADAQK 422
 Qy 421 ALEAALGKAGQOQGGILNALGQIASAAV 447
 Db 423 ALEAALGKAGQOQGGILNALGQIASAAV 449

RESULT 10

US-10-289-762-775
 ; Sequence 775, Application US/10289762
 ; Publication No. US20040006218A1

```

; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptide, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 775
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-775

Query Match      31.4%; Score 1000; DB 15; Length 212;
Best Local Similarity 99.5%; Pred. No. 2.1e-65;
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 448 VSAGVPPAAASSIGSSVQLYKTSKSTGSDYKQISAGYDAYKSIINDAYGRANDATRDV 507
Db 9 VSAGVPPAAASSIGSSVQLYKTSKSTGSDYKQISAGYDAYKSIINDAYGRANDATRDV 68
QY 508 INNVSTPALTRSPRARTARGPEKTDQALARVISGNSRTGLGVYQVSALQSVMQIIQS 567
Db 69 INNVSTPALTRSPRARTARGPEKTDQALARVISGNSRTGLGVYQVSALQSVMQIIQS 128
QY 568 NPOANNEERQKLTSAVTKPPQFGYPVQLSNDSTQKFIKLESFAEGSRTAAEIKALS 627
Db 129 NPOANNEERQKLTSAVTKPPQFGYPVQLSNDSTQKFIKLESFAEGSRTAAEIKALS 188
QY 628 FETNSLFIQQLVNLGSLYSGYLQ 651
Db 189 FETNSLFIQQLVNLGSLYSGYLQ 212

RESULT 11
US-09-841-132-436
; Sequence 436, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Rasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 436
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-436

Query Match      28.2%; Score 899.5; DB 9; Length 647;
Best Local Similarity 35.1%; Pred. No. 2.6e-57;
Matches 234; Conservative 118; Mismatches 262; Indels 53; Gaps 16;

QY 2 VNTPGPIDETETPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTDVSVERWSIL 61
Db 15 MNPIINGQI-----ASNSETKESKEA-----SPSASSSVSSWSFL 52
QY 62 RSNALMSLADKLGIASSNSSSTSRs-ADVDSSTATATPTPPPTFDYKQTAQTAQYDT 120
Db 53 SSAKHALISLRD--AILNKNSSPTDLSQLEASTSTSTVTRVAARDYNEAKSNFTAKSG 110
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEE--TATAAEWETKNADAVKGAQITELAK 179
Db 111 LENATTLAEYETKMAIDLMAALQDMERLAKQAEVTRIKAEALQEKQ-----EVIDKLNQVVK 166

```

```

QY 180 YASDNOAILDSGLKLTSPDLQALLOQVANNKKAELLKEMQDNPPVPGKTPATAQSILV 239
Db 167 LEKQNTKLTETTTTDSADQIPAINQSLEIKNSADQIIKLEGQNI---SYEAVLTNAG 223
QY 240 DQTDATATQIEKQGNARIDAVFAGQVAGNAGVENAKSNNSISNIDSAKAAIATAKTAQIAEA 299
Db 224 EVIKASSAGIKLQALQSIIVDAGQSQAAVLOAQONNSPDNIATKLLIDAEIKVNEL 283
QY 300 QKK---FPDSPILOAEQVVIQAEKDLKXIKPADGSDVPNPGTTVGGKQKQSSSGSI-- 354
Db 284 KQEHGLTDSPLVKKAEQISQAQKDIQEIKPS--GSDIPIVGPS--GSAASAGSAGVGLKS 341
QY 355 -----RVSMILDDAENETASILMSFQMIHMFENTENPDSQAQQLAAQAAK---AA 406
Db 342 SNNSGRISLLDDVDNEMAAIAMQGRSMIEQFNVNNPATAKELQAMEAQLTAMSDQLVG 401
QY 407 GDSAAALADAKALEAALGKAGQOQGLNALGQIASAAVVSAGVPPAAASSIGSSVKQ 466
Db 402 ADGEI.PABIQAIKDALAQL-KQPTDGLATAMGQVAFAAAKVGGSGAGTAGTGVQMNVKQ 460
QY 467 LYKT--SKSTGSDYKQISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRAR 524
Db 461 LYKTAFTSSTSSSYAAALSDDGYSAVKTLNLSYSESRs-GVQSAISQTANPALSRVSRSRG 519
QY 525 TEARG-PEKTDQALARVISGNSRTGLGVYQVSALQSVMQIIQSPOANNEERQKLTSA 583
Db 520 IESQGRSADASQRAAETIVRDSQTGLGVYSRLQVLDLSLSTIVSNPQVNOEIMQKLTAS 579
QY 584 VTKPQFGYPVQLSNDSTQKFIKLESFAEGSRTAAEIKALSFTETNSLFIQQLVNLG 643
Db 580 ISKAPQFGYPVQNSADSLQKFAAQLEREFVDGERSLAESRENAFRKQPAFIQQLVNLTA 639
QY 644 SLVSGYL 650
Db 640 SLFSGYL 646

RESULT 12
US-09-841-260-139
; Sequence 139, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 139
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-841-260-139

Query Match      28.1%; Score 894; DB 10; Length 660;
Best Local Similarity 34.6%; Pred. No. 6.8e-57;
Matches 233; Conservative 126; Mismatches 262; Indels 52; Gaps 17;

```



```
QY 471 --SKSTGSDYKTIQISAGYDAYKSIINDAYGRANDTRDVINNVSTPALTRSVPRARTEAR 528
Db 179 AFSSTSSSYAALSDGYKATLNSLYGERS-GVQSAISQOTANPALSRVSRSGLIESQ 237
QY 529 G-PEKTDQALARVISGNSRTLDGVYSQVSAQVQIIOGNPOANNEEIROKLTSAVTKP 587
Db 238 GRSADASQAAETIVDSQTLGVDVSRQLQVLDLSMTSTIVSNPOANQOEIMQKLTASISKA 297
QY 588 PFGPYVQVLSNDSTOKFTAKLESIFAEGRSTAAEIKALSFTFNSLFIQOVLVNI:GSLYS 647
Db 298 PFGPYPAVQNSVDSLOKFAQLEREFVQERSLAESQENAFKQPAFIOQVLVNI:ASLES 357
QY 648 GYL 650
Db 358 GYL 360

RESULT 15
US-09-971-536-69
; Sequence 69, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Bloksberg, Leonard
; APPLICANT: Lubbers, Mark
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul
; APPLICANT: Reid, Julian
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Methods
; FILE OF INVENTION: Using Them
; FILE REFERENCE: 1043C2
; CURRENT APPLICATION NUMBER: US/09/971,536
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/NZ01/00160
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 1463
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-971-536-69

Query Match 7.9%; Score 251; DB 9; Length 1463;
Best Local Similarity 22.1%; Pred. No. 3.3e-09;
Matches 159; Conservative 106; Mismatches 268; Indels 186; Gaps 27;

QY 19 ADLSAQGLEAS-----AANKSBAQRIAGBAKPKESKT--DSVERKSL-----RSAY 65
Db 720 ADVTASQAKATSAASDAASAYASEAQSIAGSHDNMEIKSLASDAEKQSIQIALAASKSAA 779
QY 66 NALMSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFFDDYKTOAQTA----- 117
Db 780 ASSSAAASAAIVASASASASASAAAVSNADASA-NSAAAYDYSEASEASASANDSSG 838
QY 118 YDT--IFTSTSLADIQALVSLQDAV-TNIKDTAATDEETAI--AAEWETKNADAVKGA 172
Db 839 YATASFAASAAAAMSAALSTAQVAKVAVSDAAAAGSAAVAVASAAQSDSKVKQATAATA 898
QY 173 -----QITELAKVASDNQAILDSLGLTSTFDLQALLOSVAN-----NNKA 214
Db 899 RSQALDDLINKIKSLTVDYASGASSASEAGQASTATSAYASASASSASEAGSYAHQAGSSA 958
QY 215 AELLKEMQDNFVPGKTPAIA-----QSLVDQDQDATATQIEKDGNAIRDAYFAG- 263
```

```
Db 959 SDAVQV--SGSAAQHASTAASAASSVYPKDSQIGTQISLAQSAASEAAKASNAASATSAAYGF 1017
QY 264 QWASGAVENAKSNNSISNIDSAXAAIATAKTQIAEPAQKFPDPSPIQBAEQKVIQAEKDL 323
Db 1018 SAASDASQAKTAASADVAVASSAASASTANSNAASAAASATKAGDS----- 1060
QY 324 KNIKPADG--SDVPNPGTIVGSKQOGSSIGSIRVSMLLDDAENETASILYSGFRQMIHM 381
Db 1061 ---KAAAGFSSAAASAAASAKGAEAVASEAASAAAS---DDSVASSAASAAAGFDK--- 1110
QY 382 FNTENPDQAAQQLAAQAARAAKAGDDSAALADAAQKALEAALGKAGQOQOQIILNALQ 441
Db 1111 -----AASRAEGRAASRAASNAASA-----AAQGTGGRSSSSASEAGQ 1148
QY 442 IASAAVVSAGVPPAAASSIGSVKQL-YKTSKSTG-----SDY 478
Db 1149 ASTATSVYASAASSASEAGSYAHQAGSSASEATGHASSATSQASAAASAKRYPDSGI 1208
QY 479 KTOISAGYDAYKSIINDAYGRANDT-----TRDVNNVSTPALTRSV 520
Db 1209 QSDVSIASSAASTASSAASAAQSEASTASSAASHASEQASIASSEDVSSSSAASVASSAA 1268
QY 521 PRARTEARGPEKTDQALARVISGNSRTLDGVYSQ-VSALQSYMQIIQS-----NPQANN 573
Db 1269 SAASSNAK-----AGNSSAAG-IYSHAASAAASAKSAESQASAAASAAASD 1314
QY 574 EIROKLTSAVTKPPQPGYPVQVLSNDSTOKFTAKLESIFAEGRSTAAEIKALSFTNS 632
Db 1315 DSVASSAASAA-----LSDDAKASSAADV-----ASSATTAATISSATSLADQS 1357
```

Search completed: March 24, 2004, 06:16:57

Job time : 153.132 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:39:25 ; Search time 11.9698 seconds
(without alignments)
2831.945 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MVNPIGPGPIDERTPPAD.....SLFIQQLVNLGSLYGYL 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	210	6.6	1120	1 STFR_ECOLI	P76072 escherichia
2	187.5	5.9	2541	1 TLN1_HUMAN	Q9Y490 homo sapien
3	186.5	5.9	1306	1 MSB2_YEAST	P23334 saccharomyc
4	186	5.8	8545	1 ANCI_CAEEL	Q9N4M4 caenorhabdi
5	182.5	5.7	1238	1 SBCC_RHOCA	O68032 rhodobacter
6	182.5	5.7	2541	1 TLN1_MOUSE	P26039 mus musculu
7	178	5.6	641	1 HRPK_PRESY	P41501 pseudomonas
8	177	5.6	2090	1 N214_HUMAN	P33588 homo sapien
9	176.5	5.5	545	1 HTR5_HALN1	Q48318 halobacteri
10	174.5	5.5	1140	1 YM96_YEAST	Q04893 saccharomyc
11	174	5.5	774	1 STF_LAMBD	P03764 bacterioph
12	172.5	5.4	3644	1 MINT_MOUSE	O62504 mus musculu
13	169	5.3	1528	1 SPAA_STRDO	P21979 streptococ
14	168.5	5.3	881	1 PRV2_YEAST	P47033 saccharomyc
15	168	5.3	810	1 HTR4_HALSA	Q48317 halobacteri
16	167.5	5.3	2492	1 TALA_DICDI	P54633 dictyosteli
17	167	5.2	810	1 HTR4_HALN1	Q9hp84 halobacteri
18	165.5	5.2	1562	1 SPAP_STRMU	P23504 streptococ
19	165.5	5.2	1555	1 PAC_STRMU	P11557 streptococ
20	163.5	5.1	2542	1 TLN2_HUMAN	Q9Y496 homo sapien
21	163	5.1	705	1 CWBA_BACSU	Q02113 bacillus su
22	160.5	5.0	1690	1 C190_DROME	O9VJ65 drosophila
23	160	5.0	535	1 HTR1_HALN1	P33741 halobacteri
24	159.5	5.0	1411	1 TCOF_HUMAN	Q13428 homo sapien
25	158.5	5.0	1379	1 YFF9_SCHPO	Q14066 schizosacch
26	158	5.0	758	1 IMMT_HUMAN	Q16951 homo sapien
27	158	5.0	763	1 HTR2_HALN1	Q9hp81 halobacteri
28	158	5.0	764	1 HTR2_HALSA	P71410 halobacteri
29	157.5	4.9	1011	1 SCA4_RICAF	Q9AJ83 rickettsia
30	157.5	4.9	1093	1 TMF1_HUMAN	P82094 homo sapien
31	157.5	4.9	1509	1 MYSN_ACACA	P05659 acanthamoeb
32	157.5	4.9	1849	1 IGMA_HAEIN	P45386 haemophilus
33	157.5	4.9	3164	1 TEGU_HSV11	P10220 herpes simp

34 157.5 4.9 5171 1 BPEA_HUMAN 094833 homo sapien
35 157 4.9 962 1 IF2_NEIMB Q9JYD2 neisseria m
36 156.5 4.9 1337 1 DEXT_STRDO P39653 streptococ
37 156 4.9 2843 1 APC_HUMAN P25054 homo sapien
38 155.5 4.9 1022 1 SCA4_RICCN Q52658 rickettsia
39 155 4.9 1161 1 DAN4_YEAST P47179 saccharomyc
40 155 4.9 1957 1 SPOF_SCHPO Q10411 schizosacch
41 154.5 4.8 757 1 IMMT_MOUSE Q8CAQ8 mus musculu
42 154.5 4.8 962 1 IF2_NEIMA Q9JTB5 neisseria m
43 154 4.8 626 1 CIB1_ECOLI P04479 escherichia
44 153.5 4.8 991 1 SCA4_RICSI Q9AJ77 rickettsia
45 153.5 4.8 2779 1 LVA_DROME Q8MS1 drosophila

ALIGNMENTS

RESULT 1

STFR_ECOLI
ID STFR_ECOLI STANDARD; PRT; 1120 AA.
AC P76072; P77560;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Side tail fiber protein homolog from lambdoid prophage Rac.
GN STFR OR B1372.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97251357; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
Science 277:1453-1474 (1997).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,
Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-Kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
DNA Res. 3:363-377 (1996).
CC -!- SIMILARITY: Belongs to the tail fiber family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----
CC EMBL; A5000234; AAC74454.1; ALT_INIT.
CC EMBL; D30774; BAAL4966.1; -
CC EMBL; D30775; BAAL4975.1; -
CC PIR; G64887; G64887.
CC EcoGene; EG13370; stfr.
CC InterPro; IPR008969; CarboxypepD_reg.
CC InterPro; IPR005003; Phage_fiber.
CC InterPro; IPR005068; Phage_fiber_2.
CC Pfam; PF03335; Phage_fiber; 6.
CC Pfam; PF03406; Phage_fiber_2; 1.


```

KW Hypothetical protein; Fiber protein; Repeat; Complete proteome.
SQ SEQUENCE 1120 AA; 113779 MW; 542E59D71EE795B4 CRC64;

Query Match 6.6%; Score 210; DB 1; Length 1120;
Best Local Similarity 20.4%; Pred. No. 0.005;
Matches 154; Conservative 102; Mismatches 294; Indels 206; Gaps 25;

QY 8 GPIDETERTPPA---DLSAQGLASAAANKSAEQAQRIAGAEAKPKESKTDVSVERWSILRS 63
DB 91 GAMEDDDARPALRRFELM---VEEVARNASAVQNTAAAKSKASDASTSAREATTAAD 147
QY 64 AVNALMSLADKLGFIASSNSSSTSRSDVDSTTATPTPPPTDDYKXQACTAYDTFT 123
DB 148 AADSARAASSTAGQAAS-SAQSASSAGTASTKATEASKSAAAAESSKSAATSGAAGT 206
QY 124 S--TSLADIQAALVSLQDAVINIKDTATDDETAIAEWETKNADAVKVGQITELAKYA 181
DB 207 SETNASASLQSAATSAATATTKASE-AATSARDAAASKERAASKE-----TNASSSA 257
QY 182 SDNQAILDSLGLTSLFDLLQALLOSVAANNKAAELLKEMQDNVPVPGKTPALIAQSLVDQ 241
DB 258 SSAASSATAAGN-----SAAKATSETNARSSETAGQASAAAGSKTAAASASNAS 310
QY 242 TDATATQIEKGNAIRDAYFQONAGSAGVENAKSNNSISNTSDSKAAIATATKQIABQAK 301
DB 311 TSA-----GQASASATAAGKSA-----ESAASSASTATTKAGEATE 346
QY 302 KFPDPSILQAEQVQIQEKDKNIKPADGSDVPNPGTTVGSGQKQSGSISIRVSMLLD 361
DB 347 Q-----ASAARSASAAKATSETNAKASE-----TSAESSKTAASASSASNAS 388
QY 362 DAENETASILMSGFRQMIMENTENPDPSQAQQLAQAARAAKAGDPSAAAAALADAQKA 421
DB 389 -----SAS-----SASAKDQATRAQSAAKSSATTASTKATEAAGSA 425
QY 422 LEALGKAGQGGIILNALGQTASAAVWSAGVPPPAASIGSVKQLYKTSKTSQSDYKQ 481
DB 426 TAAQSKSTAESAAATRAETAARABDIASAVALEDASTTKKGIVOL-----SSATNSTSETL 482
QY 482 ISAGYDAVKSNDAYGRANDATRD-----VINNVSTPALT-----RSVPRART 525
DB 483 AATP-----KAVKSAVDNAEKILQKQNGADIPDKGCFLLNNINAVSKTDFADKRGMYRVV 538
QY 526 EA-----RGPEKTDQALARY-IGNSRSLTGD----- 550
DB 539 NAPAGATSGKYVYVVMRSAGSVSELASRVITTTATRTAGDPMNCFNGFVMPGGWTR 598
QY 551 ---VTSQVSALQSVWQIICSNPOAN-NBEIRQKLTSAVTKPKQFGYPVQLS----- 598
DB 599 GRAYGMFWQYQNNERAIHSIMNSKNGDRLRSVFYVYDGAAPVFVAFIEDGLSISAPGADL 658
QY 599 --NDSTQKFLA-----KLESFLPAEG---SRT 619
DB 659 VVNDTTYKFGATNPATECIAADVILDFKSGRGFYESHSLIYVNDNLSCKLFATDDEIVARG 718
QY 620 AEIKALSFETNSLFIQO-----VIVNIGSLYSYG 649
DB 7.9 GNOIRMIGGEYALWENDGAKTYLLITNQGDVYGGW 754

RESULT 2
TLN1_HUMAN STANDARD; PRT; 2541 AA.
AC Q9Y490; Q9NZQ2; Q9UHH8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Talin 1.
GN TLN1 OR TLN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC
NCBI_TaxID=9606;

```

```

[1] SEQUENCE FROM N.A.
RP RP
RA Mao L., Fan Y.H.;
RT "Complete cDNA sequence of human talin.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079169; PubMed=10610730;
RA Ben-Yosef T., Francomano C.A.;
RT "Characterization of the human talin (TLN) gene: genomic structure,
chromosomal localization, and expression pattern.";
RL Genomics 62:316-319(1999).
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal
structures to the plasma membrane. High molecular weight
cytoskeletal protein concentrated at regions of cell-substratum
contact and, in lymphocytes, at cell-cell contacts (By
similarity).
CC -!- SUBUNIT: Bands with high affinity to vinculin and with low
affinity to integrins (By similarity).
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -!- SIMILARITY: Contains 1 I/LWEQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; AF078828; AAD13152.1; .
EMBL; AF177198; AAF23322.1; .
EMBL; AF178534; AAF27330.1; .
EMBL; AF178081; AAF27330.1; JOINED.
Genew; HGNC:11845; TLN1.
MIN; 186745; .
GO; GO:0005925; C:focal adhesion; NAS.
GO; GO:0005200; F:structural constituent of cytoskeleton; NAS.
GO; GO:0006928; P:cell motility; NAS.
GO; GO:0007016; P:cytoskeletal anchoring; NAS.
InterPro; IPR000299; Band 4.1.
InterPro; IPR002558; ILWEQ.
Pfam; PF00373; Band 41; 1.
Pfam; PF01608; I_LWEQ; 1.
ProDom; PD011820; ILWEQ; 1.
SMART; SM00295; B41; 1.
SMART; SM00307; ILWEQ; 1.
PROSITE; PS00660; FERM_1; 1.
PROSITE; PS00661; FERM_2; 1.
PROSITE; PS00057; FERM_3; 1.
PROSITE; PS50945; I_LWEQ; 1.
PROSITE; PS50945; I_LWEQ; 1.
Structural protein; Cytoskeleton.
DOMAIN 86 403
PERM.
FT DOMAIN 2340 2533
I/LWEQ.
FT CONFLICT 824 824
S -> R (IN REF. 2; AAF27330).
FT CONFLICT 1227 1227
S -> L (IN REF. 1).
FT CONFLICT 1549 1549
P -> A (IN REF. 2; AAF27330).
FT CONFLICT 1604 1604
Q -> K (IN REF. 2; AAF27330).
FT CONFLICT 1701 1701
E -> Q (IN REF. 2; AAF27330).
FT CONFLICT 1718 1718
H -> N (IN REF. 2; AAF27330).
FT CONFLICT 1966 1966
A -> R (IN REF. 1).
FT CONFLICT 2256 2256
MISSING (IN REF. 2; AAF27330).
SQ SEQUENCE 2541 AA; 269717 MW; 47FD9B22BCF47296 CRC64;

Query Match 5.9%; Score 187.5; DB 1; Length 2541;
Best Local Similarity 20.0%; Pred. No. 0.17;
Matches 157; Conservative 121; Mismatches 319; Indels 187; Gaps 32;

QY 1 MVNFIGPIDEITERTPPADLSAQGLE-----ASAANKSAEQAQRIAGAEAKPKESKTDVS 55
DB 722 VVATPISFVQEQEQLVEAGRLVAKAVGCVSQAATEDGQLLRGVGAA----- 771
QY 56 ERWSILRSVAVNALMSLADKLGFIASSNSSSTSRSDVDSTTATPTPPPTDDYKQAK 115

```



```

FT DOMAIN 5203 5274 COILED COIL (POTENTIAL).
FT DOMAIN 5339 5409 COILED COIL (POTENTIAL).
FT DOMAIN 5444 5560 COILED COIL (POTENTIAL).
FT DOMAIN 5638 5696 COILED COIL (POTENTIAL).
FT DOMAIN 5759 5879 COILED COIL (POTENTIAL).
FT DOMAIN 5938 6055 COILED COIL (POTENTIAL).
FT DOMAIN 6106 6177 COILED COIL (POTENTIAL).
FT DOMAIN 6242 6312 COILED COIL (POTENTIAL).
FT DOMAIN 6347 6463 COILED COIL (POTENTIAL).
FT DOMAIN 6541 6599 COILED COIL (POTENTIAL).
FT DOMAIN 6662 6782 COILED COIL (POTENTIAL).
FT DOMAIN 6841 6958 COILED COIL (POTENTIAL).
FT DOMAIN 7009 7080 COILED COIL (POTENTIAL).
FT DOMAIN 7145 7215 COILED COIL (POTENTIAL).
FT DOMAIN 7250 7366 COILED COIL (POTENTIAL).
FT DOMAIN 7444 7502 COILED COIL (POTENTIAL).
FT DOMAIN 7565 7685 COILED COIL (POTENTIAL).
FT DOMAIN 7744 7861 COILED COIL (POTENTIAL).
FT DOMAIN 7912 7983 COILED COIL (POTENTIAL).
FT DOMAIN 8048 8118 COILED COIL (POTENTIAL).
FT DOMAIN 8153 8204 COILED COIL (POTENTIAL).
FT DOMAIN 8273 8329 COILED COIL (POTENTIAL).
FT DOMAIN 8370 8390 COILED COIL (POTENTIAL).
FT DOMAIN 8486 8545 KLARSICHT.
FT DOMAIN 1655 1658 POLY-ARG.
FT DOMAIN 3021 3029 POLY-LYS.
FT DOMAIN 3924 3932 POLY-LYS.
FT DOMAIN 4878 4886 POLY-LYS.
FT DOMAIN 5781 5789 POLY-LYS.
FT DOMAIN 5684 5692 POLY-LYS.
FT DOMAIN 7587 7595 POLY-LYS.
SQ SEQUENCE 8545 AA; 956470 MW; 02A94D994BEE19E3 CRC64;

Query Match
Best Local Similarity 20.9%; Pred. No. 0.94;
Matches 155; Conservative 121; Mismatches 253; Indels 214; Gaps 34;

Qy 2 VNPFGPDIETERTPADLSAQGLEASAA-----NKSAEAGRIAGAAKPKSKTDSVER 57
Db 4456 VSGVKPDELDDKERAENYNDLLAR--LATAADVLNKGAELEQVAKADEK-SLHSDVDR 4512

Qy 58 WSIILRSVAVNALMSLADKLGIASSNSSSTSFSAVDSTATAPTEPPTDDYKTQATA 117
Db 4513 -----IVSRVLPL-----VRESDELHNNAEAVPTQYAPKAEELKEVEAA 4552

Qy 118 YDTIFTSTS-----LADIOAALVSLQDAVTNKDTA-----ATDEETAIAAEWETKNAD 166
Db 4553 KAVIANAPSSDAHVQQLQEQAVATAETLIPDLERARLWNEFLAARNIDIALIEQLQQPLD 4612

Qy 167 AVKGAQITELAKYAS-DNQAILDSLGKLTSLFDLLQALLQSVANNKKAELLKEMQNP 225
Db 4613 AVL--AOPKRSABEAAQDVENLRNSQQLSLD-----NKIANLQRISELPLDES-- 4661

Qy 226 VVPKTPAIAQSLVDQDTATATQTEKD-----GNAIRDVAFAGQNASGA----- 269
Db 4662 -----AYADVRFEDVDVDEOTRQYDDVLNDVAAELEDETLKQSASQVANEIDDISKM 4714

Qy 270 VENAKSNNSI-----SNIDSAKAAIATAKTIQI--AEAQKFP--PDSPFIQAEQWVIQA 319
Db 4715 IDSTDPEISILDTIAKSDIPALKAQINRIKORIVNADASRKHVTTDPKIAEDLNDKLAKL 4774

Qy 320 BKDLKN-IKPDAGSDVNPNGTVGSKQGSISIRVSMILLDAENETASILMSGRQFM 378
Db 4775 QTELDLDAIKTDEHD-----KEQ-----LILSLKLN-----ISQFEQ- 4806

Qy 379 IHMFNTENPDSQAQOELA-----AQABAAKAAGDSDSAAAB----- 414
Db 4807 IPLDOLKSDDLTKTAEKTEITNSLKPEAPEPLAKIQELREAKRVGDEARSAAHQIVALEX 4866

Qy 415 LADAQKALEAALGKAGQOQGIINALQIGTASAAVVSAG--VPPAAASSISGSSV-- 464
Db 4867 EADVTAKESAKKKKKKSPQEMIDELSAKVVEAKALIPKIEBAAKNENLPADDRKKA 4926

```

```

Qy 465 KQLYKTSKSTGSDYKTIQISAGVDAYKSINDA-----YGRANDATRDVNNVSTPALT- 517
Db 4927 EQIVSNLEAFVKVDVETQVSEKQDELKLNANDAIKRLGDLALDDEAKTVVPS-SVPALSE 4985

Qy 518 ---RSVPR-----ARTEARGPEKTDQALARVIGNSRTL----- 548
Db 4986 FKDIRIAPHLATLVEAVNDVPASVPSAVALRDA-AKFVSDLEKNTQKTGDDDEKRADELK 5044

Qy 549 GDVYSQVSALQSMQIISNPQANNEEIRKLTSAVTKPPQFGYPYVQLSNSTQKFIK 608
Db 5045 LDVGNVAKXVEDVSKYQNPQP-----LDVAKDDANKLKAT 5081

Qy 609 LBSL--FAGS-----RTAAEIK 624
Db 5082 VEQLTKLAESSDKIDPQVADIK 5104

RESULT 5
SBCC_RHOCA STANDARD; PRT; 1238 AA.
AC O68032;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nuclease sbcc subunit C.
GN SBCC.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OX Rhodobacteraceae; Rhodobacter.
CX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SB1003 / St Louis;
RA MEDLINE=97404404; PubMed=9256491;
RX Vileek C., Paces V., Maltsev N., Paces J., Haselkorn R., Fonstein M.;
RT "Sequence of a 189-kb segment of the chromosome of Rhodobacter
capsulatus SB1003".
RL Proc. Natl. Acad. Sci. U.S.A. 94:9384-9388(1997).
CC -!- FUNCTION: Sbcc cleaves DNA hairpin structures. These structures
can inhibit DNA replication and are intermediates in certain DNA
recombination reactions. The complex acts as a 3'->5' double
strand exonuclease that can open hairpins. It also has a 5'
single-strand endonuclease activity (By similarity).
CC -!- SUBUNIT: Heterodimer of sbcc and sbcd (By similarity).
CC -!- SIMILARITY: Belongs to the SMC family. SbcC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by, and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: AF010496; AAC16118.1;
DR PIR; T03465; T03465.
DR InterPro; IPR003439; ABC_transporter.
KW Hydrolyase; Nuclease; Exonuclease; Endonuclease; DNA replication;
KW DNA recombination; ATP-binding; Coiled coil.
FT NP BIND 37 44
FT DOMAIN 395 438 COILED COIL (POTENTIAL).
FT DOMAIN 466 487 COILED COIL (POTENTIAL).
FT DOMAIN 521 600 COILED COIL (POTENTIAL).
FT DOMAIN 724 770 COILED COIL (POTENTIAL).
FT DOMAIN 901 943 COILED COIL (POTENTIAL).
FT DOMAIN 1019 1052 COILED COIL (POTENTIAL).
SQ SEQUENCE 1238 AA; 128046 MW; 2B6BC6C63A859AE1 CRC64;

Query Match
Best Local Similarity 5.7%; Score 182.5; DB 1; Length 1238;
Matches 155; Conservative 116; Mismatches 329; Indels 177; Gaps 27;

Qy 6 GP-GPTDTERTPP-----ADLSAQGLEASANKSAFAQRI--AGAEAK 46

```

```

Db      293  GPAAQLELEAAAPLHLPWQAAEAARAEFLDAARSGAEAAASRSIAQAALATADAAAR 352
QY      47  PKESKTDVSVERNSILRSVAVNALMSLADKUGTASSNSSSTSRSD-----VDSTT 96
Db      353  ATAAGTETEEAFKFGRIWDRAALDAQAATATAAEARTAAETATARAAGAIRAEADL 412
QY      97  ATAPT-----PPPTFDYKTKTQATYDTTFTSTSLADIAQAALVSLQDVTNKKDAA 149
Db      413  AAETRAQAQAAEAARLAEAAQAPLADDDWQLRRDLADHRAACAQAQAAATAAQ--AA 470
QY      150  TDEETAIAEWETKNADAVKVGQATITELAKYASDNNQAILDSL-----GKLTSDLL 200
Db      471  QDRAQALAEARAAAARAAEQADQADGALAKAQAALQCALAPLEAAHPPTADLARIETD 530
QY      201  QAALLOSVAANNKAAELLEMQDNFVPGKTPAIQ-----SLVDQDQATATOI----- 249
Db      531  LAELTRADRDGAFAAALGAAQAQAAATAATEALARAQBSAAKQDLDRAEQTALATAPL 590
QY      250  EKDGNARDA-----YFAGQ-----NAGAVEN 272
Db      591  EQADLALSDARSLRAQLSAGSPVCGALEHPTPAEAGLAHLAEFLRADQAAARGAAQA 650
QY      273  AKSNNSISNDIAKAAIATAKTOIAEAKKFPDPSILOBAEQNVIOAEKDLKNIKPA-DG 331
Db      651  AR--DALTAQAQARATAEARGTQAEDQRRQAQTRAEARAANAAGDTQPRVRSARPLAPLFG 708
QY      332  SDVPNPGTIVGSK-----QQSSSIGSVSMML-----LDAENE-----TASILMSGPQ 377
Db      709  T--PDP-TALAAQCDRLCALQTAEAAQAEEISALRTLEAERDRERLRALLAHGTRE 765
QY      378  MIHMENENTPDSQAQQLAAQARAQAAG--DPSAAALADAQALEAALGKAGQQQGL 436
Db      766  RLAV-----QAETAQEAALAEARRTEAARRDGLALALA-----PALARAGEDPAA 813
QY      437  NALGQTASAAVSVAGVPPPAASISGVSQVLYKTSSTGSDYKTSQISAGYDAVKSNDAY 496
Db      814  PGLAERLAATVSVGAARTGLQAQAEALSALAPQLAAARDSDSETATAQAQSAQAARDRD 873
QY      497  G-----RARNDAATROVINN-----VSTPALTRS 519
Db      874  GAAWALRAERAPLLDGGPTALHRSRNFNDQRLAQAQAAADLATAQAALAAEAERAE 933
QY      520  VPRARTEARGPEKTDQA--LARVISGNSRTLGDVYSQVSLQSQVMQIISQNPQANNFEI-- 576
Db      934  TARAASEAATAQAAEAADLAAALAAEAAMSAAELALIALPPGTAERLRADLRGRDDAVTA 993
QY      577  -RKLTSVNTKPFQF-----GYPVQLSNDSTQKFIKLESIFAEGRSRTAAEIKALSSE 629
Db      994  ARSALGARKTDLOFIEDQGNF-----AEDPPEALSARLAALEAFEAERRAQEIGQLSAE 1045

```

RESULT 6

```

TLNI_MOUSE STANDARD; PRT; 2541 AA.
AC P26039; Q8VEF0;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Talin 1.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Fibroblast;
RX MEDLINE=91015390; PubMed=2120593;
RA Rees D.J.G., Ades S.A., Singer S.J., Hynes R.O.;
RT "Sequence and domain structure of talin.";
RL Nature 347:685-689(1990).
RN [2]

```

```

RP SEQUENCE OF 1603-2541 FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullahy S.J.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Griewood J., Schmutz J., Myers R.M.,
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal
CC structures to the plasma membrane. High molecular weight
CC cytoskeletal protein concentrated at regions of cell-substratum
CC contact and, in lymphocytes, at cell-cell contacts.
CC -!- SUBUNIT: Binds with high affinity to vinculin and with low
CC affinity to integrins.
CC -!- PTM: Phosphorylated.
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -!- SIMILARITY: Contains 1 I/LWEQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X56123; CA319588.1; -
DB EMBL; BC018557; AAH18557.1; -
DB PIR; S11661; S11661.
DB MGD; MGI:1099832; Tln.
DB GO; GO:0005925; C:focal adhesion; IDA.
DB InterPro; IPR000299; Band 4.1.
DB InterPro; IPR002558; ILWEQ.
DB Pfam; PF00373; Band 41; 1.
DB Pfam; PF01608; I_LWEQ; 1.
DB ProDom; PD011820; ILWEQ; 1.
DB SMART; SM00295; B41; 1.
DB SMART; SM00307; ILWEQ; 1.
DB PROSITE; PS00660; FERM_1; 1.
DB PROSITE; PS00661; FERM_2; 1.
DB PROSITE; PS00557; FERM_3; 1.
DB PROSITE; PS00945; I_LWEQ; 1.
DB Structural protein; Cytoskeleton; Phosphorylation.
FT DOMAIN 86 403 FERM.
FT VARIANT 2340 2533 I/LWEQ.
FT VARIANT 1105 1105 L -> P.
FT VARIANT 2180 2180 K -> M.
SQ SEQUENCE 2541 AA; 269832 MW; 14EF75ABE9FC2CBB CRC64;

```

Query Match 5.7%; Score 182.5; DB 1; Length 2541;
 Best Local Similarity 20.0%; Pred. No. 0.3;
 Matches 157; Conservative 118; Mismatches 322; Indels 187; Gaps 30;

```

QY 1 MNPIGPGPIDETPTPADLSAQGLE-----ASANKSAEQAQRIAGAAKPKESKTDV 55
Db 722 VVAPTSSVPQCPQVEAGRLVAKAVEGCVSQAQATDGLLRGVGA-- 771
QY 56 ERWSILRSVAVNALMSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFDYKTAQ 115

```


ID N214 HUMAN STANDARD; PRT; 2090 AA.
AC P35658;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nuclear pore complex protein Nup214 (Nucleoporin Nup214) (214 kDa
DE nucleoporin) (CAN protein).
GN NUP214 OR CAN OR CAN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=92195315; PubMed=1549122;
RA Von Lindern M., Fornerod M., Van Baal S., Jaegle M., De Wit T.,
RA Buijs A., Grosveld G.;
RT "The translocation (6;9), associated with a specific subtype of acute
RT myeloid leukemia, results in the fusion of two genes, dek and can,
RT and the expression of a chimeric, leukemia-specific dek-can mRNA.";
RL Mol. Cell. Biol. 12:1687-1697(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94151361; PubMed=8108440;
RA Kraemer D., Wozniak R.W., Blobel G., Radu A.;
RT "The human CAN protein, a putative oncogene product associated with
RT myeloid leukemogenesis, is a nuclear pore complex protein that faces
RT the cytoplasm.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:1519-1523(1994).
CC [1-] FUNCTION: May serve as a docking site in the receptor-mediated
CC import of substrates across the nuclear pore complex.
CC [1-] SUBUNIT: Homodimer. Interacts with DDX19 and NUP88.
CC [1-] SUBCELLULAR LOCATION: Nuclear pore complex. Cytoplasmic filaments.
CC [1-] TISSUE SPECIFICITY: Expressed in thymus, spleen, bone marrow,
CC kidney, brain and testis, but hardly in all other tissues or in
CC whole embryos during development.
CC [1-] DOMAIN: Contains F-G repeats.
CC [1-] PTM: Probably glycosylated as it reacts with wheat germ agglutinin
CC (WGA).
CC [1-] DISEASE: Implicated in a subset of acute myeloid leukemia (acute
CC nonlymphocytic leukemia) (AML) carrying a chromosomal
CC translocation t(6;9)(p23;q34) that results in the formation of a
CC DEK-CAN fusion gene.
CC [1-] DISEASE: Involved in some cases of acute undifferentiated leukemia
CC (AML) through a chromosomal translocation t(6;9)(q21;q34.1) that
CC involves NUP214/CAN and SET.
CC [1-] DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/CAN.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X64228; CAA45535.1; -;
CC PIR; S26058; S26058.
CC Genew; HGNC:8064; NUP214.
CC MIM; 114350; -;
CC GO; GO:0005643; C:nuclear pore; TAS.
CC GO; GO:0005215; F:transporter activity; TAS.
CC InterPro; IPR004325; Nucleoporin_FG.
CC InterPro; IPR001680; WD40.
CC Pfam; PF03093; Nucleoporin_FG; 22.
CC SMART; SM00320; WD40; 2.
CC Nuclear protein; Transport; Proto-oncogene; Chromosomal translocation;
CC Repeat; Glycoprotein.
KW DOMAIN 481 2076
FT DOMAIN 1409 2084
FT DOMAIN 1427 2085
FT 11 X 5 AA APPROXIMATE REPEATS.
FT 18 X 4 AA APPROXIMATE REPEATS.
FT 11 X 3 AA APPROXIMATE REPEATS.

FT DOMAIN 1213 2090 PRO/SER/THR-RICH.
FT DOMAIN 740 768 LEUCINE-ZIPPER 1.
FT DOMAIN 861 882 LEUCINE-ZIPPER 2.
FT SITE 812 813 BREAKPOINT.
SQ SEQUENCE 2090 AA; 213766 MW; 6DBE767FDD857F9F CRC64;
Query Match 5.6%; Score 177; DB 1; Length 2090;
Best Local Similarity 21.6%; Pred. No. 0.43;
Matches 161; Conservative 92; Mismatches 280; Indels 212; Gaps 32;
QY 43 AEAKPKESKTDSEVERSIARSVAVALMSLADKLGIASSNSS-----SSTSESAVDST 95
DB 897 SSAVPSQSSIHSPD--SDLESICNALL---KTTIESHTYKSLPKVPAKUSPMQAQURNF 950
QY 96 TATAPTTPP-----PPTFDYKYTOAYDTFTSTLADIQAALVLODAVNIKDTAAT 150
DB 951 LAKRKTPPVSTAPASLSRFAFLSQRYEDLDVESSTSSVSQSLES-EDARTSKDDEAV 1009
QY 151 DEE-----TAIAAEWETKNADAVKVG 172
DB 1010 VQAPRHAPVVRTPSIQFSLPHAAFPFAKSHLVHGSSPGVMGTSA---TSASKIIPQGA 1065
QY 173 QITELAKYASDNOAILDLSGLKTSFDLLQALLOSVANNKAAELL----- 218
DB 1066 DSTWLATKIVKHGAPSEPHISAPQLAAALRRQWASQAPAVNTLTSTLKNVPQVVV 1125
QY 219 KEMQDNFVPGKTPAIAQSLVDQTDAT-----ATQIEKDGNAIRDAYFAGON-ASGA 269
DB 1126 QELKNPFPATP--STAMGSSVPYSTAKTPHPVLTPVAANOAKQKGLNSLNSKPSGPTASGQ 1183
QY 270 VENAKSNNSISNIDSAKAAIATAKTQIAEAKKFPDPSPILOEAEQVMOIAEKDKLIKPA 329
DB 1184 LSSGDKASGTAKLETAVTSTPSAGQFS---KPFSPSGTGTFNGFIITTPS-SNPTAA 1239
QY 330 DG-----SDVNPNGTVGGSSKQSSIGSIRVSMLLDDAENETAS-----ILMS 373
DB 1240 QGATPSTKESQPDAPSSGGGSKPSYEAIPESSPPSGITTSASNTTPCEPAASSSRVPAS 1299
QY 374 GFRQMIHMFNTENPDQAAQ-----QELAAQA-----RAAKAGDGS-----AAAL 415
DB 1300 GTALSTTSSKLETPPSKGLLELPFSLAGETGLGFSGLRVGQA--DDSTKPTNKASSTSL 1357
QY 416 ADQKALEA-----ALGKAGQOQOILNALGQIASAAVSVAGVPAASISIGSS 463
DB 1358 TSTQPTKTSVPGSGFNTPAPFVIGKHTPE-----PVTSSATTTSVAPPAATSTSTA 1409
QY 464 V-KQLYKTSK-STG--SDYKTOISAGDYAKSINDAYGRARNDAT----- 504
DB 1410 VFGLPVTSSAGSSGVISFGGTSLSAGKTSF-----SFGSQOTNSTVPPSPPTTAATPL 1464
QY 505 -----RDVINNVSTPALTRSVPRAPTEARG---PEKTDQALARVISGNSRTIGDVG 552
DB 1465 PTFPTLTLFSGLLSSATTPSLPMSAGSSTREATSSALPEK-----PCDSEVSASAA 1515
QY 553 SQVSALQSVMOIILQSNPQANNEIRQKLTSAVKTPQFGYPPYQLNSDSTQKFIKLES 612
DB 1516 SLLEEQOSA-QLQAPQPTSD-----SVKKEPVLQAPV--SNSGTAASSTSLVAL 1563
QY 613 FAEG-----SRTAAEIKALSF 628
DB 1564 SAETATPATTGVPDARTAEVPPASSF 1588

RESULT 9

HTR5 HALN1
ID HTR5 HALN1 STANDARD; PRT; 545 AA.
AC Q483I8; Q9HP85;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Halobacterial tracheal inducer protein V.
GN HTR7 OR HTPV OR VNGI1759G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and


```
OS Halobacterium salinarium.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
CX NCBI_taxID=64091, 2242;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1;
RX MEDLINE=20504483; PubMed=11016950;
RA NG W.V., Kennedy S.P., Mahairs G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=H. salinarium; STRAIN=S9;
RX MEDLINE=96275896; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
RA Oesterheld D.;
RT "A family of halobacterial transducer proteins.";
RL FEMS Microbiol. Lett. 139:161-168(1996).
CC CC
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal
transduction.
CC CC
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC CC
CC -!- SIMILARITY: Contains 1 HAMP domain.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC CC
CC EMBL; AE005080; AAG19985.1; -.
DR EMBL; X95589; CAA64842.1; -.
DR PIR; E84327; E84327.
DR PIR; T46811; T46811.
DR HSSP; P02942; 1QUT.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR004090; Me_chemotaxis.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS01111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 1.
KW Transducer; Transmembrane; Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT DOMAIN 164 217 HAMP.
FT DOMAIN 236 476 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 545 AA; 57070 MW; FDD670389C2F428B CRC64;
Query Match 5.5%; Score 176.5; DB 1; Length 545;
Best Local Similarity 21.4%; Pred. No. 0.085;
Matches 98; Conservative 79; Mismatches 195; Indels 87; Gaps 17;
QY 12 ETERPPADLSAQGLEASANKSAEARIAGAEAKPKESKTSVERWSILRSVAVNALMSL 71
DB 118 EEERA-----EERAREKAEQQAER---QTAEASAKQARSRSAIEQLAADLSQ 169
QY 72 ADKLGIASSNSSSRSSADVDSTATAPTPPPPT-FDDYKTKQAQYAVDTI-----FRST 125
DB 170 ATEVG-ATLEASDGLTRVATDADNIAEIVATVNDMLTMTERTIDEIQFSTNVT 228
```

```
126 SLADIQAALVSLQDAVTNIKDT-----AATDEE-----TATAAEWETKNADAKVGAQITE 176
229 ASREATAGAKIEIQASQTVSEVQIEAAGTDDOREQLESVAEEMDSYSATVEEVAATAQS 288
177 LAKYASDNQALDLSGLKT-----SFDLQAAALQSVAANNKAAEILLKEMQDNPPVPGK 230
289 VADTAADTTDVA-TAGKQTAEDAIDAIDAVOETMOTTVANVDALEDUTTEIDD----- 340
231 TPAIAQSLVDTDTATATQIEKDGNAIRDVAFAGQNASG-----AVENAKSNNSI 279
341 IAEILSDIAEQTNMLA-LANATIEAARAGSGSGNGDGFVAVDEVKELATESQSAKDI 398
280 SN-IDSAAIAIAVAKTOIAEAKKFPDPSPILOEAEQMVIOAEKDLKNIKPADGSDVPNG 338
399 AELIEEVQSQATATTVEBI-----RVAQRVNDGAAAVEETVDAFGAVTENIQ 445
339 TTVGSQKQGGSIGSIKVSMLDDAENETASILMSGFRQIMHMENTENPDSSQAQQLAA 398
446 ETTDEVQE-----ISQAMDEQAQRSEKRVSS-----VDDIATISQATAD 484
399 QARAAGAAGDDSAALADAKALEAALGRAGQOQGIIN 437
485 RAENVSAASEBQ-TASITEVTSSLSQSLAAQADTLEDRIN 522
RESULT 10
YM96_YEAST STANDARD; PRT; 1140 AA.
AC Q04893;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-NOV-1997 (Rel. 35, last sequence update)
DE Hypothetical 113.1 kDa protein in PRE5-FET4 intergenic region.
GN YMR317W OR YMR924.09.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_taxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RL Nature 387:90-93(1997).
CC CC
CC -!- DOMAIN: Contains many Ser/Thr-rich domain and repeats.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC CC
CC EMBL; Z54141; CAA90835.1; -.
DR GenOnline; 142998; -.
DR SGB; S0004936; YMR317W.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;
Query Match 5.5%; Score 174.5; DB 1; Length 1140;
Best Local Similarity 17.3%; Pred. No. 0.27; 273; Indels 167; Gaps 17;
Matches 118; Conservative 126; Mismatches 126;
QY 79 SSNSSSSTSRSDVDSTATAPTPPP-----PTFDDYKTKQAQYAVDTIFSTIS 126
DB 23 SSTSTSTPTTSTTSTTSTKVTSPETIVSSSSSLVSSVVPETSSSSSLSDIASILS 82
```



```

RP RESULT 12
RX MINT MOUSE
RA ID MINT MOUSE STANDARD; PRT; 3644 AA.
AC Q62504; O80TN9; Q9P84; Q9QW2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE MSX2-interacting protein (SMART/HDAC1 associated repressor protein).
GN MINT OR SHARP OR KIAA0929.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 49-3644 FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR
RP LOCATION, TISSUE SPECIFICITY, DNA-BINDING, AND INTERACTION WITH MSX2.
RC TISSUE=Testis;
RX MEDLINE=99379811; PubMed=10451362;
RA Newberry E.P., Latifi T., Towler D.A.;
RT "The RRM domain of MINT, a novel msx2 binding protein, recognizes and
RT regulates the rat osteocalcin promoter.";
RL Biochemistry 38:10678-10690(1999).
RN [2]
RP SEQUENCE OF 1-112 FROM N.A.
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
RA Hashizume W., Hayashida K., Hirozane T., Hori F., Imotani K.,
RA Ishii Y., Itoh M., Kagawa I., Kawai J., Kojima Y., Kondo S., Konno H.,
RA Koya S., Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
RA Ohno M., Ohsato N., Saito R., Sakazume N., Sano H., Sasaki D.,
RA Sato K., Shibata K., Shiraki T., Tagami M., Takeda Y., Waki K.,
RA Watabiki A., Muramatsu M., Hayashizaki Y.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 69-3644 FROM N.A. (ISOFORM 2), AND VARIANTS THR-348;
RP PHE-762; PHE-773 AND LEU-933.
RC STRAIN=ICR; TISSUE=Brain;
RA Sakamoto T., Gotou T., Isagawa Y., Mimura H., Kimura K., Kawaichi M.;
RT "MINT/spen negatively regulates Notch signaling by inhibiting RBP-
RT J/Su(H) activity.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 318-578 FROM N.A.
RC TISSUE=Cochlea;
RX MEDLINE=97237053; PubMed=9119401;
RA Crozet F., El-Amraoui A., Blanchard S., Lenoir M., Ripoll C., Vago P.,
RA Hamel C., Fzames C., Levi-Acobas F., Depetris D., Mattei M.-G.,
RA Weil D., Pujol R., Petit C.;
RT "Cloning of the genes encoding two murine and human cochlear
RT unconventional type I myosins.";
RL Genomics 40:332-341(1997).
RN [5]
RP SEQUENCE OF 2598-3644 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT I1. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
RN [6]
RP TISSUE SPECIFICITY.
RX MEDLINE=226261914; PubMed=12374742;
RA Oswald F., Kostecka U., Astrahantseff K., Bourteelle S., Dillinger K.,
RA Zechner U., Ludwig L., Wilda M., Hameister H., Knoechel W., Lipray S.,
RA Schmid R.M.;
RT "SHARP is a novel component of the Notch/RBP-Jkappa signalling
RT pathway.";
RL EMBO J. 21:5417-5426(2002).
RN [7]
RP FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=22483652; PubMed=12594956;
RA Kuroda K., Han H., Tani S., Tanigaki K., Tun T., Furukawa T.,
RA Taniguchi Y., Kurooka H., Hamada Y., Toyokuni S., Honjo T.;
RT "Regulation of marginal zone B cell development by MINT, a suppressor
RT of Notch/RBP-J signaling pathway.";
RL Immunity 18:301-312(2003).
CC -!- FUNCTION: Essential corepressor protein, which probably regulates
CC different key pathways such as the Notch pathway. Negative
CC regulator of the Notch pathway via its interaction with RBPSUH,
CC which prevents the association between NOTCH1 and RBPSUH, and
CC therefore suppresses the transactivation activity of Notch
CC signaling. Blocks the differentiation of precursor B cells into
CC marginal zone B cells. Probably represses transcription via the
CC recruitment of large complexes containing histone deacetylase
CC proteins. May bind both to DNA and RNA.
CC -!- SUBUNIT: Interacts with NCOR2, HDAC1, HDAC2, RBP4, MBD3 and
CC MYALL1. Interacts with the nuclear receptors RAR and PPARdelta.
CC Interacts with RAR in absence of ligand. Bind to the steroid
CC receptor RNA coactivator SRA (by similarity). Interacts with MSX2.
CC Interacts with RBPSUH; this interaction may prevent the
CC interaction between RBPSUH and NOTCH1.
CC -!- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q62504-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q62504-2; Sequence=VSP_008564;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highly expressed in testis. Expressed at lower
CC level in brain, lung, spleen, liver and kidney. Weakly expressed
CC in cardiac and skeletal muscles and ovary. In spleen, it is
CC expressed in follicular B-cells, while it is weakly expressed in
CC marginal zone B-cells.
CC -!- DOMAIN: The RID domain mediates the interaction with nuclear
CC receptors.
CC -!- DOMAIN: The SPOC domain, which mediates the interaction with
CC NCOR2, is essential for the repressive activity (By similarity).
CC -!- SIMILARITY: Belongs to the Spen family.
CC -!- SIMILARITY: Contains 1 RID (receptor interacting) domain.
CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC -!- SIMILARITY: Contains 1 SPOC domain.
CC -!- CAUTION: Ref.4 sequence differs from that shown due to multiple
CC frameshifts and conflicts that create stop codons.
CC -!- CAUTION: Ref.5 sequence differs from that shown due to what seems
CC to be the presence of intronic sequence in the cDNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BY726481; -; NOT ANNOTATED CDS.
DR EMBL; AF156529; AAD55931.1; ALT_INIT.
DR EMBL; AB055980; BAB32786.1; -.
DR EMBL; Z78160; CAB01562.1; ALT_SEQ.
DR EMBL; AK122402; BAC65684.2; ALT_SEQ.
DR MGD; MGI:1991706; Mint.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF0076; rrm; 3.
DR SMART; SM00360; RRM; 3.
DR PROSITE; PS0102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE; PS00917; SPOC; 1.
KW Transcription regulation; Repressor; Nuclear protein; DNA-binding;
KW RNA-binding; Repeat; Coiled coil; Alternative splicing; Polymorphism.
FT DOMAIN 1 574 DNA-BINDING.
FT DOMAIN 6 81 RNA-BINDING (RRM) 1.
FT DOMAIN 336 416 RNA-BINDING (RRM) 2.

```

```

FT DOMAIN 439 514 RNA-BINDING (RRM) 3.
FT DOMAIN 518 590 RNA-BINDING (RRM) 4.
FT DOMAIN 559 575 COILED COIL (POTENTIAL).
FT DOMAIN 822 850 COILED COIL (POTENTIAL).
FT DOMAIN 1185 1206 COILED COIL (POTENTIAL).
FT DOMAIN 1509 1544 COILED COIL (POTENTIAL).
FT DOMAIN 1607 1627 COILED COIL (POTENTIAL).
FT DOMAIN 2216 2704 RID.
FT DOMAIN 3478 3644 SPOC.
FT DOMAIN 2138 2462 INTERACTION WITH MSX2.
FT DOMAIN 2706 2845 INTERACTION WITH RBPSUH.
FT DOMAIN 125 277 ARG-RICH.
FT DOMAIN 236 326 SER-RICH.
FT DOMAIN 648 721 TYR-RICH.
FT DOMAIN 702 832 ARG-RICH.
FT DOMAIN 2101 2233 ALA-RICH.
FT DOMAIN 2377 2518 PRO-RICH.
FT DOMAIN 2350 3475 PRO-RICH.
FT VARSPLIC 618 640 Missing (in isoform 2).
/FTid=vsp_008564.
FT VARIANT 348 348 I -> T.
FT VARIANT 762 762 S -> F.
FT VARIANT 773 773 S -> F.
FT VARIANT 933 933 S -> L.
FT CONFLICT 754 754 R -> G (IN REF. 3).
FT CONFLICT 1524 1524 D -> A (IN REF. 3).
FT CONFLICT 1560 1560 H -> Y (IN REF. 3).
FT CONFLICT 1570 1570 F -> L (IN REF. 3).
FT CONFLICT 1574 1574 R -> G (IN REF. 3).
FT CONFLICT 1609 1509 Q -> R (IN REF. 3).
FT CONFLICT 1659 1659 I -> V (IN REF. 3).
FT CONFLICT 1669 1669 S -> F (IN REF. 3).
FT CONFLICT 1705 1705 V -> A (IN REF. 3).
FT CONFLICT 1815 1815 A -> V (IN REF. 3).
FT CONFLICT 2097 2097 G -> A (IN REF. 3).
FT CONFLICT 2201 2202 MISSING (IN REF. 3).
FT CONFLICT 2322 2322 A -> V (IN REF. 3).
FT CONFLICT 2385 2385 P -> Q (IN REF. 3).
FT CONFLICT 2502 2502 R -> K (IN REF. 3).
FT CONFLICT 2505 2505 E -> K (IN REF. 3).
FT CONFLICT 2519 2519 D -> N (IN REF. 3).
FT CONFLICT 2554 2554 T -> S (IN REF. 3).
FT CONFLICT 2679 2688 LVSPFAGPVN -> VGEHPWARD (IN REF. 3).
FT CONFLICT 3010 3010 L -> P (IN REF. 3 AND 5).
SQ SEQUENCE 3644 AA; 398750 MW; 9C7EC49A81A7DA4A CRC64;

Query Match
Best Local Similarity 5.4%; Score 172.5; DB 1; Length 3644;
Matches 130; Conservative 111; Mismatches 314; Indels 143; Gaps 19;

QY 24 QGLEASAAKSAEQAQRIAGAPA-KPKESKTSQSVVERWSIL-RSVALNMLSLADKILGIASSN 81
Db 1559 QHLERKSESDFPFGRGLYGRQASEGANTSDSQEPVLFHSRWFMLTRMOCK-----EK 1613
QY 82 SSSSTSRADYDSTATAPTPPPFDYKQACTAYTITSTSLADIQAALVSLQDAV 141
Db 1614 EKQKPKAEKQEEETHPKTPPEAAETKEPEPK-----APVSAGLPAVITV 1661
QY 142 TNIKDTAATDEBTATAAEWETKNDADAVKVGQITELAKYASDNOAI-----L 188
Db 1662 VTPEPASSAPEKAEAEAPSPAGEKPAEPAPVSEETKLVSPEVSPVPEQPSQDVPPGE 1721
QY 189 DSLGKLTGFDLLQALLOSVANNKAEELLEMQDNFVVPKTPPAIAGSLVD----- 240
Db 1722 DSRDSQDS-----AALAPSAPQESAATDAVPCWNAEPLTPGTTVTSQVSESSVDPKPSQP 1776
QY 241 -----QTDATATOIEK-----DGNALRDAYAGQNASGAVENAKSNNSINIDSAK 286
Db 1777 LSKLTQRSEAEKGEKPEPTTPSTPEPATONAGVASEAQPPASDEVANPPVAAKDKRT 1836
QY 287 AAATATQIAEAKKFPDPSILOAE-----QMWIQAEKDLKN*325
Db 1837 NKSKEKTSVQAAASAVVEKPVTRKSERIDREKLKSSPRGEAQKLELXWAEKIRT 1896

```

```

QY 326 IKPADGSDVNPCTTGGSKQSGSIGSRVSNMLDDAENE-----TASILMSGF 375
Db 1897 ASKSSGGDTEHPSPSLPSRRNRNVSVVATWTDHESRSPAKEPVEQPRVTKRLEREL 1956
QY 376 RQMIHMFNT-----ENPDSQAQOELAAQABAKA-----AGDSSAAALADAQ 419
Db 1957 QEAVVPTTPFRGRPEKTRRAEDGEHERKEFAETTPFAEGRWSRSPRSQKSAANA 2011
QY 420 KALEAALGAGQOQILNALGQIASAAVVSAGVPPA-----AASSIGSSVKQLYKTSKSTG 475
Db 2012 -----GPGQKGRNEQKVEAAEAAGAQAQSTREGNPKSRGREGAASEPKRDERDSTDKSGP 2067
QY 476 SDYKTIQISAGYDAYKINDAYGEARNDA TRDVINNV-----STPALTRSVPRARTEARGPE 531
Db 2068 DTFPEVLEKRPPEKTYKSKRGGRAR--STRSGMDRAAHQSRLEMAAAGAAQAAKKEAGP- 2124
QY 532 KTDQALARVIGNSRTILGDVYSQVSAIQSMQIIOQNPNANNEEIRQKLTSAVTKPQFG 591
Db 2125 -----AAASPOSESPOKSGSSPQLA-----NNPADPDREAEEESASASTAPPEGT 2171
QY 592 YPVVQLSNDSTOKFIKAL-----ESLFAEGSRRTAAEIKK 625
Db 2172 QLAQQLLEQAVQNIKALPEPSAAASKGTATATATAA 2209

RESULT 13
SPAA_STRDO
ID SPAA_STRDO STANDARD; PRT; 1528 AA.
AC P21979;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell surface antigen I/II precursor.
GN SPAA.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91310320; PubMed=1855987;
RA Lapollia R.J., Haron J.A., Kelly C.G., Taylor W.R., Bohart C.,
RA Hendricks M., Pyatt J., Graff R.T., Ma J.K.-C., Lehnert T.;
RT "Sequence and structural analysis of surface protein antigen I/II
RT (Spaa) of Streptococcus sobrinus."
RL Infect. Immun. 59:2677-2685(1991).
RN [2]
RP SEQUENCE OF 423-817 FROM N.A.
RX MEDLINE=90299827; PubMed=1694526;
RA Goldschmidt R.M., Thoren-Gordon M., Curtiss R. III;
RT "Regions of the Streptococcus sobrinus spaa gene encoding major
RT determinants of antigen I."
RL J. Bacteriol. 172:3988-4001(1990).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -1- MISCELLANEOUS: IMMUNODOMINANT DETERMINANTS ARE LOCATED IN THE
CC C-TERMINAL TWO-THIRDS OF THE SPAA PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SPAP/SSP5/SPAA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X57841; CAA40973.1; -.
DR EMBL; M38210; AAA26977.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.

```

```

RESULT 14
PRY3_YEAST
ID_PRY3
AC P4703
DT 01-FEB
DT 01-FEB
DT 28-FEB
DE PRY3
GN PRY3

```

```

Db 594 DNRKHTSYGSSGASLDSLRTTTSISVSSNTT-----QLVSTCTSESYS----- 640
Qy 394 QELAAQARAAGAAGDGAALADQAALAKAGAGQOQOGLNALGQIAAAGVVSAGVP 453
Db 641 -----DPSFPAISATATT-----ESNLITWT-----ITASCSTDSNFP 673
Qy 454 PAAASS-----IGSSVKQL-----YKTSKSTGSDYKQTOISAGYDAYKSI-----ND 494
Db 674 TSAASSTDETAFTRTISTSCSTLNGASTQSELITSPMKTNVVPASSFPSTTTTCLEND 733
Qy 495 --AYGRARNDATRDVNNVSTPALTRSPVPRARTARGPEKTDQMLARVISGNSRTLGDVY 552
Db 734 DTAFSSITYEYNAATIIN---PGETSSLASDPAFSEKNEPSTSVKSTNEGST--- 785
Qy 553 SOVSALQSVQIIIOGNPOANNEUR-----CKLTSVATKPKQFQYVYVQLSND 600
Db 786 --TTYQOTVATLAKPSSTSLGARTTGSNGRSTTSQDGSAMHQPTSSYITQKEGTS 843
Qy 601 STQKPIAKLESFAGSRTAAEIKALSPETNSL-----FQQVL 639
Db 844 TTKALSA-----YEGATPLSI-----PQCNSLAGTIAAFVAVL 878

RESULT 15
HTR4 HALSA
ID HTR4 HALSA STANDARD; PRT; 810 AA.
AC Q48317;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Halobacterial transducer protein IV.
GN HTR5 OR HTRIV.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=S9;
RX MEDLINE=96275896; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
RA Oesterheld D.;
RT "A family of halobacterial transducer proteins.";
RL FEWS Microbiol. Lett. 139:161-168(1996).
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction.
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X95589; CAA64841.1; -
CC PIR; T46810; T46810.
CC HSSP; P02942; 10U7.
CC InterPro; IPR004089; Chmtaxis_transd.
CC InterPro; IPR003660; HAMP.
CC Pfam; PF00672; HAMP; 1.
CC Pfam; PF00015; MCPsSignal; 1.
CC SMART; SM00304; HAMP; 2.
CC SMART; SM00283; MA; 1.
CC PROSITE; PS50111; CHEMOTAXIS_TRANSDC_2; 1.
CC PROSITE; PS50885; HAMP; 2.
KW Transducer; Transmembrane; Repeat.
FT DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT TRANSVMEM 39 59 POTENTIAL.
FT DOMAIN 60 323 EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSMEM 324 344 POTENTIAL.
FT DOMAIN 345 810 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 345 397 HAMP 1.
FT DOMAIN 439 493 HAMP 2.
FT DOMAIN 512 748 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 810 AA; 85207 MW; 37B0F6046A39D9BA CRC84;

Query Match
Best Local Similarity 19.7%; Pred. No. 0.36;
Matches 144; Conservative 103; Mismatches 316; Indels 168; Gaps 26;

Qy 11 DETERPPADLSAQ-----GLEASANKSA-----EAORIA-----GAEA 45
Db 86 DWLERNEQSVLIASNNPRLGNTTAADKQAVYVQLVAAELDADRIADVHVADPTVVGASD 145
Qy 46 KPRESKTDSEVERSIILRSATNALMSLADKLQ-----LIASSNS-----SSSTSRSDVD 93
Db 146 ARIVASTDEARGTRVSADTHPWDRTRSGRTVTSSTNPNYRTAGGQVVSSMSVAADLT 205
Qy 94 STTATAPTPPPPTFDYKTOAQATYDTFTFTSLADIQAALVSLQDAVTNI----- 144
Db 206 HVLVVEYTA-----GDLSDQFGAGIDGTFQVVRPTSDATAVLFSDDAGTDAVGQYIPDR 260
Qy 145 --KDTAATDEETALAAEWETKNADAV-----KVGQAQITELAK-----YASDNQ 185
Db 261 SQSEIPALDSATEQGGPTNTPTKDSVLDREYVAAAYTTPGKNVWVVKHAPSESALFNSQ 320
Qy 186 AILDSLG-----KLTSPDLL-----QAALQSVANNKAAELKEMQDNPVPGKTPAI 234
Db 321 IRGILGFILVALGVVVGVTIGRTAAAVQSL--SAAAEIEAGNYDNDVASSRDEI 378
Qy 235 AQ---SLVDQDTATATQIEKGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKA-AIA 290
Db 379 GOLFASTGSMRDALVTQIDEAAREAQEAQODAE--AERAEADAREARAEADAKADAEA 436
Qy 291 TAKQTAEAKKXFPDPILOEAECQVIOAEKDLKNIKPADGSDVPNPCTVGGSKQOGSS 350
Db 437 LAELEQAQER-----YSDVMAACADGDUITRRMPA----- 466
Qy 351 IGSTRVSMILLDDAENETASTILNSGFRQMIMFNTENPDSQAQQLAAQARAARAAAGDSDS 410
Db 467 -----DDTDNEAMAAIASFNEMLAQWEHTIIDIOEFADAVATASEAEVGNAD- 515
Qy 411 AAAALADAQKALEALGKAGQOQOGLNAL-GQIA--SAAVVSAGVPPAAASSIGSSVKOL 467
Db 516 AERASGQSVESVQEIAGAAGEQRMMLDTVSGEMTDLSSAIEEVAASADSVAEHSHTAEI 575
Qy 468 YKTSKSTGSD-----YKTIISAGYDAYKSIINDAYGRAEN--DATRDVINNVSTPALTR 518
Db 576 ARGEQTAEDAIERSISVQBAIDATVQNVBALDDQWAEISEIVDLISDIAEQTNMLALNA 635
Qy 519 SVPRARTEARG-----PEKTDQALARVISGNSRTILGDYVYSQVSAALQSVQMII 565
Db 636 NIEAARADKSGDGFVAVVADSVKDLABETQESAGDI---ERRITEVQSQTATVAERAA 691
Qy 566 QSNPQAN---NEIROKLSAVTKPQFGYVYVQLSN-----DSTQKFTAKLESFPAEG 616
Db 692 EESMDAGIDAVEEVVDAFTAVSDHADETTGTGVOEISDTTDDDOASSTEEAVSMTEEVADLS 751
Qy 617 SRTAAEIKALS 627
Db 752 DSTAGEAQSVS 762

Search completed: March 24, 2004, 05:56:56
Job time : 28.9698 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:47:55 ; Search time 49.0374 Seconds
(without alignments)
4188.690 Million cell updates/sec

Title: US-10-608-559-2
Perfect score: 3187
Sequence: 1 MVNIGPDIETERTPPAD.....SLFIQQVLVNTGSLYSGVLQ 651

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_orzanelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	16 Q9Z7H7	Q9Z7H7 chlamydia p
2	2205	69.2	715	2 Q46166	Q46166 chlamydia p
3	1276.5	40.1	636	16 Q824X3	Q824X3 chlamydia p
4	899.5	28.2	647	16 Q84627	Q84627 chlamydia t
5	247	7.8	2310	16 Q8CMU7	Q8cmu7 staphylococ
6	232.5	7.3	2055	2 Q85472	Q85472 abiotrophia
7	232	7.3	1822	2 Q07290	Q07290 streptococ
8	231.5	7.3	2478	2 Q9RL69	Q9rl69 staphylococ
9	229.5	7.2	2478	2 Q9LCH2	Q9lch2 staphylococ
10	223.5	7.0	2481	16 Q99QR6	Q99qr6 staphylococ
11	221.5	7.0	973	16 Q8XQ04	Q8xdq4 escherichia
12	220	6.9	6713	16 Q99U54	Q99u54 staphylococ
13	217.5	6.8	6713	16 Q93IR6	Q93ir6 staphylococ
14	215.5	6.8	2016	5 Q9BI70	Q9bit0 pleocitreure
15	214	6.7	842	5 Q86RNL	Q86rnl plasmodium
16	214	6.7	2275	16 Q8NUJ3	Q8nuj3 staphylococ

17	214	6.7	9904	16 Q8NWQ6	Q8nwq6 staphylococ
18	213.5	6.7	3360	16 Q88XB6	Q88xb6 lactobacill
19	212	6.7	855	5 Q86DH2	Q86dh2 plasmodium
20	211	6.6	9439	16 Q8CP76	Q8cp76 staphylococ
21	210.5	6.6	1795	16 Q9LQJ9	Q9lqj9 staphylococ
22	210	6.6	2271	16 Q99QI4	Q99qv4 staphylococ
23	209	6.6	625	5 Q86RN6	Q86rn6 plasmodium
24	209	6.6	2167	2 Q845L8	Q845l8 mycoplasma
25	208.5	6.5	3072	2 Q939N5	Q939n5 streptococ
26	208.5	6.5	4776	16 Q97P71	Q97p71 streptococ
27	207.5	6.5	856	5 Q86RM6	Q86rm6 plasmodium
28	207	6.5	621	5 Q86RM2	Q86rm2 plasmodium
29	206.5	6.5	762	16 Q89IU6	Q89iue bradyrhizob
30	205	6.4	625	5 Q86RN2	Q86rn2 plasmodium
31	204	6.4	803	5 Q86RN4	Q86rn4 plasmodium
32	204	6.4	1313	5 Q9NE28	Q9ne28 leishmania
33	204	6.4	3381	2 Q9KX33	Q9kx33 streptococ
34	201.5	6.3	849	5 Q86R71	Q86r71 plasmodium
35	201	6.3	1336	16 Q9ACV2	Q9acv2 streptomyce
36	200.5	6.3	1545	16 Q9RDQ1	Q9rdq1 streptomyce
37	200	6.3	732	5 Q86RM7	Q86rm7 plasmodium
38	200	6.3	842	5 Q86RN3	Q86rn3 plasmodium
39	199	6.2	1365	2 Q49525	Q49525 mycoplasma
40	199	6.2	2178	2 Q9KWR3	Q9kwr3 streptococ
41	198.5	6.2	1932	16 Q7UQ94	Q7uq94 rhodopirell
42	198	6.2	849	5 Q86RN3	Q86rn3 plasmodium
43	198	6.2	1222	16 Q87XX8	Q87xx8 pseudomonas
44	197.5	6.2	838	5 Q86RN7	Q86rn7 plasmodium
45	197.5	6.2	2283	2 Q8VQ99	Q8vq99 staphylococ

ALIGNMENTS

RESULT 1
Q9Z7H7
ID Q9Z7H7 PRELIMINARY; PRT; 651 AA.
AC Q9Z7H7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CHLPN 76 kDa HOMOLOG 1 (C7622) (Hypothetical protein).
GN CPN0728 OR CPJ0728 OR CP0018 OR CPB0756.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CNL029;
RX MEDLINE=9206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RA "Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.";
RL Nat. Genet. 21:385-386 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Linher K., Weidman J., Kouri H., Craven B., Bowman C., Dodson R.,
RA Winn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RA "Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.";
RL Nucleic Acids Res. 28:1397-1406 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RA "Comparison of whole genome sequences of *Chlamydia pneumoniae* J138


```
RT from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-TW-183;
RA Geng M.M., Schumacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; AE001654; AAD18867.1; -
DR ENBL; AE002165; AAR37914.1; -
DR ENBL; AP002547; BAA98935.1; -
DR ENBL; AE017159; AAP98695.1; -
DR PIR; D72042; D72042.
DR PIR; E86581; E86581.
DR PHCI-2DPAGE; Q927H7; -.
DR TIGR; CP0018; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 651 AA; 68217 MW; 47AB6C3FF2FF0123 CRC64;

Query Match 100.0%; Score 3187; DB 16; Length 651;
Best Local Similarity 100.0%; Pred. No. 2.9e-137;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
DB 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
QY 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTTATATPTPPPTFDDYKTKQAQTAAYDT 120
DB 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTTATATPTPPPTFDDYKTKQAQTAAYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKY 180
QY 181 ASDNQAILDSLGKLTSPDLLQAALLQSVANNKKAELLKEMODNPVVPKTPAIAQSLVD 240
DB 181 ASDNQAILDSLGKLTSPDLLQAALLQSVANNKKAELLKEMODNPVVPKTPAIAQSLVD 240
QY 241 QTDATATQIEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
DB 241 QTDATATQIEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
QY 301 KKPPDSPILQEAQMVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQGGSSIGSIRVSMML 360
DB 301 KKPPDSPILQEAQMVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQGGSSIGSIRVSMML 360
QY 361 DDAENETASILMSGFRQMIHMENTENPDQAQOELAAQARAARAKAGDSDSAAALADAQK 420
DB 361 DDAENETASILMSGFRQMIHMENTENPDQAQOELAAQARAARAKAGDSDSAAALADAQK 420
QY 421 ALEAALGKAGQGGQILNALGQIASAAVVSAGVPPAAASSIGSSVQLYKTSKSTGSDYKT 480
DB 421 ALEAALGKAGQGGQILNALGQIASAAVVSAGVPPAAASSIGSSVQLYKTSKSTGSDYKT 480
QY 481 QISAGYDAYKSIDNAYGRANDATRVINNVSTPALTRSPVPRATEARGPEKTDQALARY 540
DB 481 QISAGYDAYKSIDNAYGRANDATRVINNVSTPALTRSPVPRATEARGPEKTDQALARY 540
QY 541 ISGNSRTLGDVYSQVSALQSVQMLIQSNPQANNEIEIKLTSAVTTPKPPQFGYPVQLSND 600
DB 541 ISGNSRTLGDVYSQVSALQSVQMLIQSNPQANNEIEIKLTSAVTTPKPPQFGYPVQLSND 600
QY 601 STOKFIAKLSLFAEGSRTRAEIKALSFETNSLFIQQVLNIGSLYSGYLQ 651
DB 601 STOKFIAKLSLFAEGSRTRAEIKALSFETNSLFIQQVLNIGSLYSGYLQ 651

RESULT 2
Q46166
ID Q46166 PRELIMINARY; PRT; 715 AA.
```

```
AC Q46166;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 76 kDa protein.
OS Chlamydia pneumoniae (Chlamydiales; Chlamydiaceae; Chlamydiales).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94156481; PubMed=7509320;
RA Perez-Melgosa M., Kuo C.-C., Campbell L.;
RT "Isolation and characterization of a gene encoding a Chlamydia
RT pneumoniae 76-kilodalton protein containing a species-specific
RT epitope.";
RL Infect. Immun. 62:880-886(1994).
DR EMBL; L23921; AAA23117.1; -.
DR PIR; I40729; I40729.
SQ SEQUENCE 715 AA; 76626 MW; 3F01C9A59DE964C5 CRC64;

Query Match 69.2%; Score 2205; DB 2; Length 715;
Best Local Similarity 99.6%; Pred. No. 1.5e-92;
Matches 452; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
DB 257 LVPPIGPGPIDETERTPPADLSAQGLEASAAANKSAEQAQRIAGAEAKPKESKTSVERWSI 316
QY 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTTATATPTPPPTFDDYKTKQAQTAAYDT 120
DB 317 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTTATATPTPPPTFDDYKTKQAQTAAYDT 376
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKY 180
DB 377 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKY 436
QY 181 ASDNQAILDSLGKLTSPDLLQAALLQSVANNKKAELLKEMODNPVVPKTPAIAQSLVD 240
DB 437 ASDNQAILDSLGKLTSPDLLQAALLQSVANNKKAELLKEMODNPVVPKTPAIAQSLVD 496
QY 241 QTDATATQIEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
DB 497 QTDATATQIEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 556
QY 301 KKPPDSPILQEAQMVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQGGSSIGSIRVSMML 360
DB 557 KKPPDSPILQEAQMVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQGGSSIGSIRVSMML 616
QY 361 DDAENETASILMSGFRQMIHMENTENPDQAQOELAAQARAARAKAGDSDSAAALADAQK 420
DB 617 DDAENETASILMSGFRQMIHMENTENPDQAQOELAAQARAARAKAGDSDSAAALADAQK 676
QY 421 ALEAALGKAGQGGQILNALGQIASAAVVSAGVPP 454
DB 677 ALEAALGKAGQGGQILNALGQIASAAVVSAGVLP 710

RESULT 3
Q824X3
ID Q824X3 PRELIMINARY; PRT; 636 AA.
AC Q824X3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN CCA00015.
OS Chlamydia caviae.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364;
```



```

Db 1798 NAVQAAQKKAIAIENDPALTEEEKDAKAKADYDAEAKAKDAIDAATSNADVTQAKDAG 1857
QY 506 DVINNV-STPA---LTRVPRPAREARPEKTDQALARVISGNSRTTLGDVYSQVSALQS 560
Db 1858 DAINAVPQTAKTDAKNAVDAQADKKSIAIENDPALTR--BEKDAVRKAVDAEAKKAKD 1915
QY 561 VMOIIQSPPQANNEIROKLTSAVTKPPQ 589
Db 1916 AIDAATSNADVTAKQTEG--TOAINAVPQ 1942

RESULT 7
Q07290 ID Q07290 PRELIMINARY; PRT; 1822 AA.
AC Q07290;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Epf* protein.
GN EPF*.
OS Streptococcus suis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1307;
RN [1]_TaxID=1307;
RP SEQUENCE FROM N.A.
RC STRAIN=1890;
RX MEDLINE=93328288; PubMed=8335363;
RA Smith H.E., Reek F.H., Vecht U., Gielkens A.L.J., Smits M.A.;
RT "Repeats in an extracellular protein of weakly pathogenic strains of
RT Streptococcus suis type 2 are absent in pathogenic strains.";
RL Infect. Immun. 61:3318-3326(1993).
DR EMBL; X71880; CAA50714.1; -
DR PIR; S33441; S33441.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos YSIRK.
DR Pfam; PF04650; YSIRK signal; 1.
DR TIGRFAMs; TIGR01168; YSIRK signal; 1.
SQ SEQUENCE 1822 AA; 192632 MW; 3838960C77641D7D CRC64;

Query Match
Best Local Similarity 21.4%; Pred. No. 0.027;
Matches 166; Conservative 127; Mismatches 289; Indels 192; Gaps 34;

QY 10 IDETERTPPADISAGOLEASANKSAEAGRIAGABAKPKESKTDV---ERWSILRSANV 66
Db 1072 IDDNPNLTPE-----KESAKNAVEEAHVATA-AIDKASTPDVAVQVEEDKGV--AAIN 1122
QY 67 ALMSLADKLGIASSNSSSTSRSD-----VDSTTAT-----APTPTPTFDYKT 112
Db 1123 LITAKADAGVIAKLAIDIKLEBKQABAEAKAIDAATNNEEKAIAKALQDVVDKGA 1182
QY 113 Q-----AQTAYDTFTSTSLADIQAALV-----SLQDAVTNIXKDTAATDETAIAAE-WETK 163
Db 1183 ELEDAAARVATNEIHEATTTEKAKAEELAGEKSLITDGTGEARDAVELAKDELAKEAIRTE 1242
QY 164 NADAVKVGQIITELAKYA-----SDNQ-----ILDSIGKLTSTPL 199
Db 1243 EEEATKIVEKLAEDTRKATIEDNPNLSDEKQAEIKKLTDAVAKTLATIRDNADKRTQ--- 1299
QY 200 LQALLOSVANNKKAELKEMQDNVPVPGKTPAIAQSLVDOTDATA-TQIEKQGNATRD 258
Db 1300 -PAEKAQALADLEKAKETQK-IADKAAIDRLILVKDGELEATKQDAKNKAKDAARAKE 1357
QY 259 AVFAGQNASGA-----VENAKSNNSIS-----NIDSAKAAIA----- 290
Db 1358 AIASPNLTDAEKKTFTDAVDAEAVAKANDAIASAATSPADVQKEEDAGVAAIAEDVLDAAK 1417
QY 291 -TAKTQIAE-----AQKFPDPSILOBAEQMVTQAEKDLKNIKPAGDSVPNPGTIVGGSK 345
Db 1418 QDAKNKIADAAAARAKEAIGSNPNLTDAEKKTFTDAVDAEAVAKANDAIASAATSPADV--QK 1475

```

```

QY 346 QGSSIGSIRVSM---LDDAENETASILMSGFROMIHMFNTENPDSSAAQOELAAQARA 402
Db 1476 BEDAGVAAIEDVDLDAAKQDAKNTAKESDAKSAIDANPNLTDAEKESAKKAVDADAKA 1535
QY 403 AKAGDDSSAAALADAQKALEAALGKAGQ-----QQGILNALGQIASAA----- 446
Db 1536 ATDAID--ASTSPVEAQSAEDKGVSGIAQDVLDAAKQAKNKIAKEVAAAKEAIDANPNL 1593
QY 447 -----VVSAGVPPAAASS-----IGSSVKQLYKTSKSTGSDYKT 480
Db 1594 SDAEKEASKKAVDADAKATTDAIDASTSPVEAQSAEDKGVSGIRQDVLDAAK---QDAKN 1650
QY 481 QISAGYDAYKSINDAYGR---ARNDATRDVIN---NVSTPALTRVPRPAREARPE--KT 533
Db 1651 KIAKESDAKSAIDANPNLTDAEKESAKKAVDADAKAATDAIDAST--SPVEAQSAEDKG 1708
QY 534 DOALARVISGNSRTTLGDVYSQVS-ALQSVMOIIQSNP-----CANNEEIROKLTSAVT 585
Db 1709 VGATAKQILDAAKQ--DAKNKIAKEAESAKSVISNPNLTDAAKEAAKSEIDKAVEEAIV 1766
QY 586 -----KPPQFGYPYVQVLSNDSTQKPIAKLESIFAPGS 617
Db 1767 LINGVRTYQELEKIKLPMAALIKPAAKVTPVD--PNLTKELIARIKAFLEKENN 1819

RESULT 8
Q9RL69 ID Q9RL69 PRELIMINARY; PRT; 2478 AA.
AC Q9RL69;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Mip protein.
GN MRP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]_TaxID=1280;
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=97302526; PubMed=9158773;
RA Wu S., de Lencastre H., Sali A., Tomasz A.;
RT "A phosphoglucotomase-like gene essential for the optimal expression
RT of methicillin resistance in Staphylococcus aureus: molecular cloning
RT and DNA sequencing.";
RL Microb. Drug Resist. 2:277-286(1996).
RN [2]
RC SEQUENCE FROM N.A.
RX MEDLINE=97431478; PubMed=9286983;
RA Jolly L., Wu S.W., Van Heijenoort J., de Lencastre H.,
RA Mengin-Lecreulx D., Tomasz A.;
RT "The femR315 gene from Staphylococcus aureus, the interruption of
RT which results in reduced methicillin resistance, encodes a
RT phosphoglucosamine mutase.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=99265121; PubMed=10332717;
RA Wu S., de Lencastre H.;
RT "Mip-a new auxiliary gene essential for optimal expression of
RT methicillin resistance in Staphylococcus aureus.";
RL Microb. Drug Resist. 5:9-18(1999).
DR EMBL; Y09927; CAB55329.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos YSIRK.
DR Pfam; PF04650; YSIRK signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; YSIRK signal; 1.
SQ SEQUENCE 2478 AA; 263031 MW; 6B9659A02D023C74 CRC64;

```

```
Query Match      7.3%; Score 231.5; DB 2; Length 2478;
Best Local Similarity 22.8%; Pred. No. 0.042;
Matches 169; Conservative 128; Mismatches 285; Indels 159; Gaps 35;

Qy 4 PIGPGPIDE-----TERPPADLSAQGLEASANKSA-EACRIAGABAKPKESKIDS--- 54
Db 780 PLNPDTTNEEVAERINAAKVS--GVKAIEATTTAQDLERVKNEEISKIENTIDTSQT 837

Qy 55 -----VERWSILRSVAVNALMS-----LADKLG-----IASSNSSSSTS 87
Db 838 KMDAYNEVKQATARKAQNATVSATNNEEVAEDAAVDAQAQKGLHDIQVVKSKQEVADT 897

Qy 88 RSADVSTTA-TAPTPPPPTDDYKTAQATYDT-----IFTSTSLADIQAALVSL--- 137
Db 898 KSKVLDKINALQIQAKVKPAAD--TEVENAYNTRKQEIQNSNASTTEEKQAAVTELDTK 954

Qy 138 -QDAVNIKDTAATD-----BETAIAAEWETKNADAVKGAQITELAKYASDNOAILDS 190
Db 955 KQEARTNL-DAANTNSDVTTAKDNSIAAINQVQAATTKSDAK-AEIAQKASERKTAIEA 1012

Qy 191 LGKLTSPDLLQAA---LQSVANNKAAELLKEMQDNVPVPGKT----- 231
Db 1013 MNDSTT-EEQQAADKVDQAVVTAN--ADIDNAAANDVDNKTNEATIAATPDANVK 1069

Qy 232 PAIAQSLVDQDATATQIE-KDGNAIRDAYFAGQ-----NAGAVENAKSNSI---- 279
Db 1070 PAAQAIADKVOAQETALDGNNGSTTEEKAQAQVQVTEKTTADAIDAHTNAEVEAAK 1129

Qy 280 -----SNIDSAKAAIAT---AKTOIAEAKKFPDPSPIQLEA--EQMVIQA 319
Db 1130 KAATAKIEAIQPAATTTKDNKAEIATKANERKTAIAQTDITABEIAAANADVNAVTOA 1189

Qy 320 EKDLNLIKPADG-SDVPNPGTTVGSGKQOGSSIGSVKLYKTSKTSQSDYKTOISAGYDAYK 378
Db 1190 N---SNIEAANSQNDVQAKTGTGENSIDQVTPVVKAT-----ARNEITAILNNKQEI 1241

Qy 379 IHMNTENPDSQAQQLAQAARAKAAGDSDSAAA-----ALADAQKALEALGKAGQ 431
Db 1242 QATPDATDEEKQADAE--ANTENGKANQAIASAATTNAQVDEKANAEEAINAVTPKVVK 1299

Qy 432 QOGLILNALGQI-ASAAVVSAGVPPAAASSIGSSVKOLYKTSKTSQSDYKTOISAGYDAYK 490
Db 1300 KQAAKDEIDQIQATQTNVINNDQNATTEEKEAAIQQL-----ATAVTDKANNITATD--- 1352

Qy 491 SINDAYGRANDTRDVINNVSVPALTRSPRARTPEARPEKPTDQALARVIGNSRSLTGD 550
Db 1353 --DNGVQDAK-DAGKNSIQSTQ-----PATAVKSNKNDVDQAVTTQONQIDNTTGA 1401

Qy 551 VYSQVSALQSVMQIIQSNPOANNEIROKLTSAVTKPPQFGYPYVQ-LSNDSTOKFIACL 609
Db 1402 TTEKNAAKDL--VLKAKAYQDILNAQTINDVTQIKDOAVADIQGITADTTIKDVAK- 1458

Qy 610 ESLFAEGSRTAAEIKALSFT 630
Db 1459 ----DELATKANEQKALIAQT 1475

RESULT 9
Q9LCH2
ID Q9LCH2 PRELIMINARY; PRT; 2478 AA.
AC Q9LCH2;
DT 01-OCT-2000 (TreeBrel. 15, Created)
DT 01-OCT-2000 (TreeBrel. 15, Last sequence update)
DT 01-JUN-2003 (TreeBrel. 24, Last annotation update)
DE FmtB.
GN FmtB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=20348625; PubMed=10896508;
```

RESULT 10

```

Q99QR6 PRELIMINARY; PRT; 2481 AA.
AC Q99QR6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FmtB protein.
GN FmtB (MRP) OR FmtB OR SAV2160 OR SAI1964.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacilliales; Staphylococcus.
OX NCBI_TaxID=158978, 158979;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
RA Kanchisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58322.1; -.
DR EMBL; AP003136; BAB43253.1; -.
DR PIR; D90011; D90011.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016774; F:phosphotransferase activity, carboxyl group. .; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0016310; P:phosphorylation; IEA.
DR InterPro; IPR000890; Acetate kin.
DR InterPro; IPR005877; Gpos YSIRK.
DR Pfam; PF04650; YSIRK_signal; I.
DR TIGRfam; TIGR01167; LPXTG_anchor; I.
DR TIGRfam; TIGR01168; YSIRK_signal; I.
DR PROSITE; PS01075; ACETATE_KINASE_1; 1.
KW Complete proteome.
SQ SEQUENCE 2481 AA; 263765 MW; ELEAB99B81665E8 CRC64;

Query Match 7.0%; Score 223.5; DB 16; Length 2481;
Best Local Similarity 22.1%; Pred. No. 0.097;
Matches 163; Conservative 129; Mismatches 293; Indels 153; Gaps 31;

QY 4 PIGPGPIDE-----TERTPADLSAQGLEPASAANKSA-EAQRAGAEAKFKESKTDSDS--- 54
DB 763 FLNPFTTNEEVAELERINAQVSV--GVKAIETATTAQLERVKNEIFKINENITDSTQT 820
QY 55 -----VERWSILRSVAVNMLSLADKLGIASNSSSSSTSRSD-----VDSTTTATAPT 101
DB 821 KMDAYKEVQQAATARKAQAATVSNATDEBVAEANAADAAQTEGLHDIQVVKSQQEVADT 880
QY 102 PP-----PPTFDYKTOQAQVADT-----IFTSTSLADIQAALVSL--- 137
DB 881 KAKVLDKINAIQTAQVKAAPAD---TEVENAYNTRKQEIQNASTTEKEAAYTDLAK 937
QY 138 -QDAVTNKDITATD-----EETALAEWEKTNADAVKGAQITELAKYASDNCAILDS 190
DB 938 QKEATNL-DAANTSVDVTTAKDNGIAINQVAATTKSKDAK-AEIAQKASERTALIEA 995
QY 191 LKGLT-----SFDLQALLOSVAANNKA-----AELKEMQDNFVVPFGKT--PAI 234
DB 996 MNDSTTEEQAAKQKVDQAVVTANADIDNATANTDNDNAKTNTNEATIAAITPDANVKPA 1055
QY 235 AQLVDDQDTATQIE-KDGNRAIRDAYFAGQN-----ASGAVNAKNSNI----- 279
DB 1056 KOAIADKQVQOETAIDANNSTTEKEAKQVQTEKTAADAADAIDAAHSNVEVEAAKNAE 1115

```

```

QY 280 -----SNIDSAAKAIAT-----AKTOIAEAKKPPSPILQEA--EOMVIOAEKD 322
DB 1116 IAKIEAIOFATTTKDNKAQAIATKANERKTAIAQTODITAEIAAANADVDNAVQTAN-- 1173
QY 323 LKNIRPADG-SDVPNPGTTVGSGKQGSIGSIRVSMLLDDAENETASTILMSGFPMIHM 381
DB 1174 -SNIEAANSQNDVDQAKTIGTSETSDQVTFVNVKKAT-----ARNBITAILNKLKEIQAT 1227
QY 382 FNTENPDQAAQOELAAQARAAGDSDSAA-----ALADAQKALEALGKAGQOQO 434
DB 1228 PDATDEEQADAE--ANTENGKANQAIASAATTNAQVDEAKANAEAAINAVTPKVVKKQA 1285
QY 435 LINALGOI-ASNAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKTOISAGYDAYKXIN 493
DB 1286 AKDEIDQLQATQTNVINNDQATNEEKAIIQOL-----ATAVTDKANNITA----- 1332
QY 494 DAYGGRANDVARDVINNVSTPALTSVPRARTEARGPPEKTDQALARVISGNSRSLTGDVVS 553
DB 1333 -----ATDNGVDVTAADAGKNSIQSTQATATVAKSNAKNEVDQAVTTQQAIDNTTGATE 1387
QY 554 QVSALQSVMIQSNPQANNEIEIKQLTSVATKPPQFGYVYQ-LSNDSTOKFIKLESL 612
DB 1388 EKNAAKDL--VLKAKEKAYQDILNAQTNDVTVQIKDQAVADIQGITADTTIKDVAK--- 1441
QY 613 FAEGSRTRAEIKALSFET 630
DB 1442 -DELATKANEOKALIAQT 1458

RESULT 11
Q8XDQ4 PRELIMINARY; PRT; 973 AA.
AC Q8XDQ4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative membrane protein of prophage CP-933X (Putative tail fiber
DE protein).
GN 21918 OR BCS1650.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller J.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Arantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Lida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005333; AAG56007.1; -.
DR EMBL; AP002555; BAB35073.1; ALT_INIT.
DR PIR; B90835; B90835.
DR PIR; C85693; C85693.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR008969; Carboxypeptidase.
DR InterPro; IPR005003; Phage_fiber.

```


Db 1737 ESSAATEIYQETQIASSIAAGALGEAHTSELAEEASSSSSAASAAAAEQSLYDTSSA 1796
QY 555 VSAIQSV-----MGIQSNPANNNEEROKLTSAVTTPPGFGYVYVLSNDSTQK 604
Db 1797 ASASASSDFIASDITNQSLSVNSASSAAEEVSQVDETYQ---NFDQYSSISASA 1853
QY 605 FIAKLESFAEGSRVAAETK---ALSFTNSLFIQQVLVNIIGSLYS 647
Db 1854 SAAQSSEIYQDVSSSSAAASTSSAASLETSGTVAESGSGTAASSSYA 1900
RESULT 15
Q86RN1 PRELIMINARY; PRT; 842 AA.
ID Q86RN1
AC Q86RN1;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Merozoite surface protein 3 alpha (fragment).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22346636; PubMed=12458823;
RA Rayner J.C., Corredor V., Feldman D., Ingravallo P., Iderabdullah F.,
RA Galinski M.R., Barnwell J.W.;
RT "Extensive polymorphism in the plasmodium vivax merozoite surface coat
protein MSP-3alpha is limited to specific domains."
RL Parasitology 125:393-405(2002).
DR EMBL; AF491951; AAC20882.1; -.
FT NON TER 842
SQ SEQUENCE 842 AA; 89463 MW; E0B13FA7D9C65169 CRC64;

Query Match 6.7%; Score 214; DB 5; Length 842;
Best Local Similarity 19.9%; Pred. No. 0.066;
Matches 137; Conservative 107; Mismatches 291; Indels 154; Gaps 22;
QY 27 EASAANKSAEAPORAG-----AEAKPKESKTDVSVERWSILRSAY--NA 67
Db 163 QAAIAEKADAABEAGAKENKLDVQSVKIAEEASTKAKDKTEAEIAVEIVKAVVAKEE 222
QY 68 LMSLADKLGTASNSSSSRSDVDSSTATATPTPPPTDDYKTQAOTHTYDIFTSTSL 127
Db 223 AQKASDEAQACEKAKAHAKAKASDTTKTET-----FKTNAEAAKAKAKAGN 274
QY 128 ADIQAAALVSLQDAVTNFKDTAATDETA-----JAEWETKNADAVKVGQITELAKY 180
Db 275 ANKAAATEAESANELSVAKQKADAEAAKAEKQVKAETAAEVAKAKVAKAEAEAKK 334
QY 181 ASDNQAILDSGLKLTSPDLLQAALLOSVAANNKAAELLKEMQDNVVPVPGTPTAIAQSLVD 240
Db 335 ABEAKKIVDKIAQDTKVPEQAQAEFATETVKATTAATEAGKNAQAEKLPAAEAT--- 391
QY 241 QTDATATQIKDGNATDAYFAGQASGAVENAKSNNSISNIDSAKAAIATATQIAEQ 300
Db 392 -----SDAVKGKADAAEAKAAGEAKKASITETI-----AIEVAKAEVLNAE 431
QY 301 KKPDPSPILQEAQEWIQAELKLNKIPA-----DGSDPNPNPGTTVGSKQSQSSIG 352
Db 432 VK-----KTAQAEAKDTEAEQAQAKAAAEAKTHGEKAEKVESTKAHSDKAQENKK 487
QY 353 SIRVSMILLDDAENETASILMSGFPMIHFNFNTENPDSQA----- 391
Db 488 AKEAS---DEAENRAVDALBEAYAVEAHLARTKNAEBSAKSATDLSKLEBAKEAIDAAN 544
QY 392 -AQELAAQAPAAKAGDSDSAALA--DAQALAEALGKAGQOQGIINALGQIASAAV 448
Db 545 IAHQKWLKATQAATTIAKEKKAQVAAEKTQKEATAAKLKAQAEK-KGAETEAVKAAVE 603
QY 449 SAGVPPAA---ASSIGSSVKLYKT-----SKSTGSDYKTQISAGVDAYKINSIDAYGR 498

Db 604 ARAAAEAKQEAQKVGAS-KEPQETKKNKANTVEAEATGNKAK-----KAEDAAEEVKERAKK 658
QY 499 ARNDATRDVINNVSTPALTRSVPRARTEARPEKTDQALARVIGNSRTGLGVYSQVSAL 558
Db 659 A-NEATD---ANVARSEADKAIAAAKAKAKAREKAAAYGLLKTKN----- 698
QY 559 QSVMOIIOQSNPQ-----ANNEEIRQKLTSAVTTPPGFGYVYVLSNDSTQKFIKLESL 612
Db 699 QYVLEPLDISPADNITSKEEQVKEVED-----QODEDNSE---AEVEEA 742
QY 613 FAEGSRTAAEIKALSFTNSLFIQQVLVN 641
Db 743 LPNGSGAQEEDVNLEMDDEEVEENVATN 771

Search completed: March 24, 2004, 05:59:09
Job time : 61.0374 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:04 ; Search time 713.142 Seconds
(without alignments)
11032.393 Million cell updates/sec

Title: US-10-608-559-3

Perfect score: 1852
Sequence: 1 atgagtgcagataagct.....gatcttcaataataaaga 1852

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

1: geneseq1980s.*

2: geneseq1990s.*

3: geneseq2000s.*

4: geneseq2001as.*

5: geneseq2001bs.*

6: geneseq2002s.*

7: geneseq2003as.*

8: geneseq2003bs.*

9: geneseq2003cs.*

10: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1852	100.0	1852	3 AAD02064	Aad02064 5'-trunca
2	1848.8	99.8	2156	3 AAD02063	Aad02063 Chlamydia
3	1836.2	99.1	110000	2 AAX91990_08	Continuation (9 of
4	1748.8	94.4	1956	4 AAS57031	Aas57031 C. pneumo
5	1748.8	94.4	1956	6 ABL92612	Ab192612 Chlamydia
6	1748.8	94.4	1956	6 ABL91190	Ab191190 Chlamydia
7	1748.8	94.4	1956	9 ADD42825	Add42825 Chlamydia
8	1161.8	62.7	2238	3 AAD02066	Aad02066 C. pneumo
9	1151	62.1	1456	3 AAD02065	Aad02065 3'-trunca
10	187.6	10.1	110000	2 AAZ01425_07	Continuation (8 of
11	187.6	10.1	110000	2 AAZ01425_08	Continuation (9 of
12	187.2	10.1	1537	4 AAS63292	Aas63292 C. tracho
13	187.2	10.1	1537	4 AAS6196	Aas6196 Chlamydia
14	187.2	10.1	1537	6 ABL92425	Ab192425 Chlamydia
15	186	10.0	1944	6 ADD43867	Add43867 Chlamydia
16	185.4	10.0	1941	6 ABL92619	Ab192619 Chlamydia
17	184	9.9	1171	4 AAS56981	Aas56981 C. tracho
18	184	9.9	1171	9 ADD42775	Add42775 Chlamydia
19	184	9.9	1834	4 AAS57006	Aas57006 C. tracho
20	184	9.9	1834	9 ADD42800	Add42800 Chlamydia
21	183	9.9	1983	4 AAS56996	Aas56996 C. tracho
22	183	9.9	1983	9 ADD42790	Add42790 Chlamydia
23	148.4	8.0	150	2 AAV16207	Aav16207 Part of t

24	51.6	2.8	2205	8	ADA28999	Ada28999 DNA encod
25	49.8	2.7	2000	7	ADA71938	Ada71938 Rice gene
26	49	2.6	896	2	AAQ68903	Aaq68903 PspA prot
27	49	2.6	946	2	AAX33124	Aax33124 Streptoco
28	49	2.6	957	2	AAT61726	Aat61726 Streptoco
29	49	2.6	1860	7	ACA49648	Aca49648 Prokaryot
30	49	2.6	1990	2	AAT61725	Aat61725 Streptoco
31	49	2.6	1990	7	ABX95373	Abx95373 S. pneumo
32	49	2.6	2085	2	AAQ78131	Aaq78131 Pneumococ
33	49	2.6	2085	2	AAT08979	Aat08979 DNA encod
34	49	2.6	2085	2	AAT07178	Aat07178 Pneumococ
35	49	2.6	2085	2	AAV33264	Aav33264 Pneumococ
36	49	2.6	2085	2	AAV39470	Aav39470 Streptoco
37	49	2.6	2085	2	AAZ25063	Aaz25063 Streptoco
38	49	2.6	2085	2	AAV84069	Aav84069 DNA encod
39	49	2.6	2085	2	AAV84069	Aav84069 DNA encod
40	47.4	2.6	956	7	ABX95374	Abx95374 S. pneumo
41	47.4	2.6	2085	2	AAQ28674	Aaq28674 PspA gene
42	47.2	2.5	110000	6	ABA90521_07	Continuation (8 of
43	47	2.5	9373	6	ABL33291	Ab133291 Human imm
44	47	2.5	9373	6	ABK31317	Abk31317 Signal tr
45	47	2.5	9373	6	ABL70560	Ab170560 Chemicall

ALIGNMENTS

RESULT 1
AAD02064
ID AAD02064 standard; DNA; 1852 BP.
XX
AC AAD02064;
XX
DT 26-MAR-2001 (first entry)
XX
DE 5'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.
XX
KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
KW upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; treatment; truncation mutant; ds.
XX
OS Chlamydia pneumoniae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1752
FT /*tag= a
FT /product= "5'-truncated Chlamydia pneumoniae 76KDa
FT protein"
FT /transl_except= (pos:1489..1491, aa:lle)

XX
WO200066739-A2.
09-NOV-2000.
03-MAY-2000; 2000WO-CA000511.
03-MAY-1999; 99US-0132270P.
30-JUN-1999; 99US-0141276P.
(AVET) AVENTIS PASTEUR LTD.
Murdin AD, Oomen RP, Wang J, Dunn P;
WPI; 2000-687542/67.
P-PSDB; AAY71955.
Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
for vaccinating against Chlamydia infections.
Claim 2b; Page 97-99; 112pp; English.

CC The present sequence is a DNA coding for 5'-truncated Chlamydia
CC pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the
CC diagnosis, prevention and treatment of C. pneumoniae infections (e.g.
CC pneumonia, upper respiratory tract disease, bronchitis, sinusitis and
CC acute respiratory disease such as cough, sore throat, hoarseness, fever;
CC and abnormal chest sounds on auscultation). C. pneumoniae sequence is
CC also used as vaccines for immunising humans against diseases caused by C.
CC pneumoniae
XX

SQ Sequence 1852 BP; 578 A; 409 C; 407 G; 458 T; 0 U; 0 Other;

Query Match 100.0%; Score 1852; DB 3; Length 1852;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAGTCGGCAGATPAAGCTGGTATTGCTTCTAGTAACAGCTGCTCTTACTAGCAGA	60
DB	1	ATGAGTCGGCAGATPAAGCTGGTATTGCTTCTAGTAACAGCTGCTCTTACTAGCAGA	60
QY	61	TCTGACAGCTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCAGCTTTGAT	120
DB	61	TCTGACAGCTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCAGCTTTGAT	120
QY	121	GATTATAGACTCRAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACTAGCT	180
DB	121	GATTATAGACTCRAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACTAGCT	180
QY	181	GACATACAGGCTGCTTTGGTAGCCTCCAGAGTGTCTCACTAATATAAGATACAGCG	240
DB	181	GACATACAGGCTGCTTTGGTAGCCTCCAGAGTGTCTCACTAATATAAGATACAGCG	240
QY	241	GCTACTGATAGGAAACCGCAATCGCTCGGAGTGGGAAATPAAGATGCCGATTCCT	300
DB	241	GCTACTGATAGGAAACCGCAATCGCTCGGAGTGGGAAATPAAGATGCCGATTCCT	300
QY	301	AAAGTTGGCGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGGATCTT	360
DB	301	AAAGTTGGCGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGGATCTT	360
QY	361	GACTCTTTAGGTAAAGTCACTGCTTCCGACCTCTTACAGGCTGCTCTTCTCCATCTGTA	420
DB	361	GACTCTTTAGGTAAAGTCACTGCTTCCGACCTCTTACAGGCTGCTCTTCTCCATCTGTA	420
QY	421	GCAACAATAACAAAGCGAGCTGCTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	480
DB	421	GCAACAATAACAAAGCGAGCTGCTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	480
QY	481	GGGAAACGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG	540
DB	481	GGGAAACGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG	540
QY	541	ATAGAGAAGATGGAAATGCGATTAGGATGATATTTTGCAGGACAGAAACGCTAGTGA	600
DB	541	ATAGAGAAGATGGAAATGCGATTAGGATGATATTTTGCAGGACAGAAACGCTAGTGA	600
QY	601	GCTGTAGAAATGCTAAATCTTAATACAGATATAGCAACATAGATTCAGCTAAAGCAGA	660
DB	601	GCTGTAGAAATGCTAAATCTTAATACAGATATAGCAACATAGATTCAGCTAAAGCAGA	660
QY	661	ATCGCTACTGCTTAAGACCAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAATT	720
DB	661	ATCGCTACTGCTTAAGACCAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAATT	720
QY	721	CTTCAAGAGCGGAAACAAATGGTAAACAGGCTGAGAAAGATCTTAAAAAATATCAAACT	780
DB	721	CTTCAAGAGCGGAAACAAATGGTAAACAGGCTGAGAAAGATCTTAAAAAATATCAAACT	780
QY	781	GCAGATGGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACAGGA	840
DB	781	GCAGATGGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACAGGA	840
QY	841	AGTACTATTGGTAGTATTCGTTCCATGCTGTTAGATGATGCTGAAATGAGACGCT	900
DB	841	AGTACTATTGGTAGTATTCGTTCCATGCTGTTAGATGATGCTGAAATGAGACGCT	900

RESULT 2

AAD02063

ID AAD02063 standard; DNA; 2156 BP.

XX

AC

XX

DT 15-SEP-2003 (revised)
 DT 26-MAR-2001 (first entry)
 DE Chlamydia pneumoniae 76 kDa full-length protein encoding DNA.
 DE 76 kDa protein; bactericidal; diagnosis; prevention; treatment;
 KW pneumonia; upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; ds.
 XX Chlamydia pneumoniae.
 OS
 PH Location/Qualifiers
 FT 101..2056
 FT /*tag= a
 FT /product= "Chlamydia pneumoniae 76kDa protein"
 PN WO200066739-A2.
 XX
 XX 09-NOV-2000.
 PD
 PP 03-MAY-2000; 2000WO-CA000511.
 PP
 XX 03-MAY-1999; 99US-0132270P.
 PR 30-JUN-1999; 99US-0141276P.
 PR
 XX (AVET) AVENTIS PASTEUR LTD.
 PA
 XX Murdin AD, Oomen RP, Wang J, Dunn P;
 PI WPI; 2000-687542/67.
 DR P-PSDB; AAY71954.
 DR
 XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
 PT for vaccinating against Chlamydia infections.
 XX
 XX Claim 2a; Fig 1; 112pp; English.
 PS
 CC The present sequence is a DNA coding for Chlamydia pneumoniae 76 kDa
 CC protein. C. pneumoniae 76 kDa protein is used in the diagnosis,
 CC prevention and treatment of C. pneumoniae infections (e.g. pneumonia,
 CC upper respiratory tract disease, bronchitis, sinusitis and acute
 CC respiratory disease such as cough, sore throat, hoarseness, fever; and
 CC abnormal chest sounds on auscultation). C. pneumoniae sequence is also
 CC used as vaccines for immunising humans against diseases caused by C.
 CC pneumoniae. (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 2156 BP; 672 A; 461 C; 471 G; 552 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1848.8; DB 3; Length 2156;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 60
 DB 305 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 364
 QY 61 TCTGCAGACGTGGACTCAACGACGACGACCGACCTAGCTCCCTCCACCCACGTTTGTAT 120
 DB 365 TCTGCAGACGTGGACTCAACGACGACGACCGACCTAGCTCCCTCCACCCACGTTTGTAT 424
 QY 121 GATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT 180
 DB 425 GATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT 484
 QY 181 GACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGCTCAATAATAAGGATACAGCG 240
 DB 485 GACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGCTCAATAATAAGGATACAGCG 544
 QY 241 GCTACTGATGAGGAACCGCAATCGCTGGGAGTGGGAAACTAAGAAATGCCGATGCACTT 300
 DB 545 GCTACTGATGAGGAACCGCAATCGCTGGGAGTGGGAAACTAAGAAATGCCGATGCACTT 604

QY 301 AAAGTTGGCGCGCAAAATTAAGAAATAGCGAAATATGCTTCGGATATACCAAGCGATTCTT 360
 DB 605 AAAGTTGGCGCGCAAAATTAAGAAATAGCGAAATATGCTTCGGATATACCAAGCGATTCTT 664
 QY 361 GACTCTTTAGTAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTGA 420
 DB 665 GACTCTTTAGTAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTGA 724
 QY 421 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
 DB 725 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 784
 QY 481 GGGAAAACGCTGCAATGTCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
 DB 785 GGGAAAACGCTGCAATGTCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 844
 QY 541 ATAGAGAAAGATGGAATGGGATGCGATTTAGGATGCAATTTTTCAGGACAGAAAGCTAGTGA 600
 DB 845 ATAGAGAAAGATGGAATGGGATGCGATTTAGGATGCAATTTTTCAGGACAGAAAGCTAGTGA 904
 QY 601 GCTGTAGAAAATGCTAAATCTTAATAACAGTATAAGCAACATAGATTGAGCTTAAAGCAGCA 660
 DB 905 GCTGTAGAAAATGCTAAATCTTAATAACAGTATAAGCAACATAGATTGAGCTTAAAGCAGCA 964
 QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCCAAAT 720
 DB 965 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCCAAAT 1024
 QY 721 CTTCAAGAAGCGGAACAAATGGTAAATACAGGCTCAGAAAAGATCTTAAAAATATCAAACT 780
 DB 1025 CTTCAAGAAGCGGAACAAATGGTAAATACAGGCTCAGAAAAGATCTTAAAAATATCAAACT 1084
 QY 781 GCAGATGGTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGGA 840
 DB 1085 GCAGATGGTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGGA 1144
 QY 841 AGTAGTATGTAGTATTTGCTGTTTCCATGCTGTGTAGATGATGCTGAAAATGAGACCGCT 900
 DB 1145 AGTAGTATGTAGTATTTGCTGTTTCCATGCTGTGTAGATGATGCTGAAAATGAGACCGCT 1204
 QY 901 TCCATTTTGTATGCTGGGTTTTCGTGATGATTCACATGTTTCAATACGGAATCTCTGAT 960
 DB 1205 TCCATTTTGTATGCTGGGTTTTCGTGATGATTCACATGTTTCAATACGGAATCTCTGAT 1264
 QY 961 TCTCAAGTGGCCAAACAGGAGCTCGAGCAACAGCTAGACAGCAAGCGCTGGAGAT 1020
 DB 1265 TCTCAAGTGGCCAAACAGGAGCTCGAGCAACAGCTAGACAGCAAGCGCTGGAGAT 1324
 QY 1021 GACAGTGTCTGCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAGCGGCTCTAGGTAAA 1080
 DB 1325 GACAGTGTCTGCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAGCGGCTCTAGGTAAA 1384
 QY 1081 GCTGGCAACAAACAGGCGATCACTCAATGCTTTAGGACAGATGCTTCTGCTGTGTTGTG 1140
 DB 1385 GCTGGCAACAAACAGGCGATCACTCAATGCTTTAGGACAGATGCTTCTGCTGTGTTGTG 1444
 QY 1141 AGCGCAGAGTTTCTCCCGCTGCAGCAAGTTCTATAGGTTGATCTGTAAACAGCTTTAC 1200
 DB 1445 AGCGCAGAGTTTCTCCCGCTGCAGCAAGTTCTATAGGTTGATCTGTAAACAGCTTTAC 1504
 QY 1201 AAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGCAGGTTTATGATGCT 1260
 DB 1505 AAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGCAGGTTTATGATGCT 1564
 QY 1261 TACAAATCCATCAATGATGCTATGGTAGGCGACGAAATGATGCGACTCGTGATGTGATA 1320
 DB 1565 TACAAATCCATCAATGATGCTATGGTAGGCGACGAAATGATGCGACTCGTGATGTGATA 1624
 QY 1321 AACAAATGTAGTACCCCGCTCTCAGACGATCGTTCCTAGAGCACGAAACAGAGCTCGA 1380
 DB 1625 AACAAATGTAGTACCCCGCTCTCAGACGATCGTTCCTAGAGCACGAAACAGAGCTCGA 1684
 QY 1381 GGACCAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGCAATAGCAGAACTCTT 1440

Db	1685	GGACAGAAAAAAGAGATCAAGCCCTCGCTAGGGTGAATTTCTGGCAATAGCAGAACTCTT	1744
Qy	1441	GGAGATGCTATAGTCAAGATTTTCGGCACTACAACTCTGTAATGAGATCACTCAGTCGAAT	1500
Db	1745	GGAGATGCTATAGTCAAGATTTTCGGCACTACAACTCTGTAATGAGATCACTCAGTCGAAT	1804
Qy	1501	CCTCAGGGAATATAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGCCCTCCA	1560
Db	1805	CCTCAGGGAATATAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGCCCTCCA	1864
Qy	1561	CAGTTTGGCTATCCTTATGTGCAACTTCTTAATGACTCTACACAGAAAGTTTCATAGCTAAA	1620
Db	1865	CAGTTTGGCTATCCTTATGTGCAACTTCTTAATGACTCTACACAGAAAGTTTCATAGCTAAA	1924
Qy	1621	TTAGAAAGTTTGTGCTGAGGATCTAGGACAGCAGCTGAAATAAAGACACCTTCCTTT	1680
Db	1925	TTAGAAAGTTTGTGCTGAGGATCTAGGACAGCAGCTGAAATAAAGACACCTTCCTTT	1984
Qy	1681	GAACGAACTCCTTGTATTCAGAGGTGCTGTCATATATCGGCTCTCTATATCTTGGT	1740
Db	1985	GAACGAACTCCTTGTATTCAGAGGTGCTGTCATATATCGGCTCTCTATATCTTGGT	2044
Qy	1741	TATCTCCAATAACACACCTTAAGTGTTCGTTGGAGAGATTATATGCTGCTTTCGTAAGG	1800
Db	2045	TATCTCCAATAACACACCTTAAGTGTTCGTTGGAGAGATTATATGCTGCTTTCGTAAGG	2104
Qy	1801	CCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATAAATAAAGA 1852	
Db	2105	CCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATAAATAAAGA 2156	
RESULT 3			
AAx91990_08/c			
Continuation (9 of 13) of AAx91990 from base 800001 (Nucleotide sequence of the complete			
WP Sequence split into 13 fragments Locus AAx91990 Accession Aax91990			
WP	Fragment Name	Begin	End
WP	AAx91990_00	1	110000
WP	AAx91990_01	100001	210000
WP	AAx91990_02	200001	310000
WP	AAx91990_03	300001	410000
WP	AAx91990_04	400001	510000
WP	AAx91990_05	500001	600000
WP	AAx91990_06	600001	710000
WP	AAx91990_07	700001	810000
WP	AAx91990_08	800001	910000
WP	AAx91990_09	900001	1010000
WP	AAx91990_10	1000001	1110000
WP	AAx91990_11	1100001	1210000
WP	AAx91990_12	1200001	1230025
Query Match 99.1%; Score 1836.2; DB 2; Length 110000;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1849; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			
Qy	1	ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60	
Db	28394	ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 28335	
Qy	61	TCTCGACAGTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTTTGGAT 120	
Db	28334	TCTCGACAGTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCAGTTTGGAT 28275	
Qy	121	GATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACTCAACATCACTAGCT 180	
Db	28274	GATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACTCAACATCACTAGCT 28215	
Qy	181	GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACTAATAAAGGATACAGCG 240	
Db	28214	GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACTAATAAAGGATACAGCG 28155	
Qy	241	GCTACTGATAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGATGCGGATGCGAGTT 300	

Db	28154	GCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGATGCCGATGCAGTT 28095	
Qy	301	AAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGGATCTTT 360	
Db	28094	AAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGGATCTTT 28035	
Qy	361	GACTCTTTAGGTAAACGTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA 420	
Db	28034	GACTCTTTAGGTAAACGTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA 27975	
Qy	421	GCAAAACAATAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGTCCCA 480	
Db	27974	GCAAAACAATAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGTCCCA 27915	
Qy	481	GGGAAAACGCTCCAAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540	
Db	27914	GGGAAAACGCTCCAAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 27855	
Qy	541	ATAGAGAAAGATGGAATGGGATTAGGGATGCATATTTTGCAGGACAGAACGCTAGTGA 600	
Db	27854	ATAGAGAAAGATGGAATGGGATTAGGGATGCATATTTTGCAGGACAGAACGCTAGTGA 27795	
Qy	601	GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATGATTTCAGCTAAAGCAGCA 660	
Db	27794	GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATGATTTCAGCTAAAGCAGCA 27735	
Qy	661	ATCGCTACTGCTAAGACACAATAATGCTGAAGCTCAGAAAAGTTCCCGGACTCTCCAAT 720	
Db	27734	ATCGCTACTGCTAAGACACAATAATGCTGAAGCTCAGAAAAGTTCCCGGACTCTCCAAT 27675	
Qy	721	CTTCAAGAAAGCGGAAACAATGTTAATACAGGCTGAGAAAGATCTTAAATAATATCAACCT 780	
Db	27674	CTTCAAGAAAGCGGAAACAATGTTAATACAGGCTGAGAAAGATCTTAAATAATATCAACCT 27615	
Qy	781	GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCCAAGCAACAAGGA 840	
Db	27614	GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCCAAGCAACAAGGA 27555	
Qy	841	AGTAGTATTTGGTAGTATTCGTGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT 900	
Db	27554	AGTAGTATTTGGTAGTATTCGTGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT 27495	
Qy	901	TCCATTTTGTGATGCTGGGTTTCGTGATGATTCACATGTTCAATACGGAATAATCTGAT 960	
Db	27494	TCCATTTTGTGATGCTGGGTTTCGTGATGATTCACATGTTCAATACGGAATAATCTGAT 27435	
Qy	961	TCTCAAGCTGCCCAACAGGAGCTCGCAGCACAAAGCTAGACGACGAAAGCCGCTGGAGAT 1020	
Db	27434	TCTCAAGCTGCCCAACAGGAGCTCGCAGCACAAAGCTAGACGACGAAAGCCGCTGGAGAT 27375	
Qy	1021	GACAGTGTCTGCTGAGGCTGGCAGATGCTCAGAAAGCTTTAGAAAGGCTCTAGGTAA 1080	
Db	27374	GACAGTGTCTGCTGAGGCTGGCAGATGCTCAGAAAGCTTTAGAAAGGCTCTAGGTAA 27315	
Qy	1081	GCTGGGGCAACAACAGGGCATCTCAATGCTTTAGGACAGATCGTCTCTGCTGCTG-TTGT 1139	
Db	27314	GCTGGGGCAACAACAGGGCATCTCAATGCTTTAGGACAGATCGTCTCTGCTGCTGTTGT 27255	
Qy	1140	GAGCGCAGGAGTTTCCTCCCGCTGACGAAAGTTCTATAGGGTCACTGTGTAACACAGCTTTA 1199	
Db	27254	GAGCGCAGGAGTTTCCTCCCGCTGACGAAAGTTCTATAGGGTCACTGTGTAACACAGCTTTA 27195	
Qy	1200	CAGAGCTCAAAATCTACAGGTTCTGATTTAATAACACAGATATCAGCAGGTTATGATGC 1259	
Db	27194	CAGAGCTCAAAATCTACAGGTTCTGATTTAATAACACAGATATCAGCAGGTTATGATGC 27135	
Qy	1260	TTACAAATCCATCAATGATGCCCTATGGTAGGGCAACGAAATGATGCCACTCGTGTGAT 1319	
Db	27134	TTACAAATCCATCAATGATGCCCTATGGTAGGGCAACGAAATGATGCCACTCGTGTGAT 27075	
Qy	1320	AAACAATGTAAGTACCCCGCTCTCAAGATCCCGTTCTTAGAGCAGCAACAGAGCTCG 1379	
Db	27074	AAACAATGTAAGTACCCCGCTCTCAAGATCCCGTTCTTAGAGCAGCAACAGAGCTCG 27015	

QY 1380 AGGACAGCAAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCT 1439
 Db 27014 AGGACAGCAAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCT 26955
 QY 1440 TGGAGATGCTATAGTCAAGTTTGGGCACTACAACTCTGAATGCGAGTCACTCAGTCGAA 1499
 Db 26954 TGGAGATGCTATAGTCAAGTTTGGGCACTACAACTCTGAATGCGAGTCACTCAGTCGAA 26895
 QY 1500 TCTCAAGCGAATAATAGGAGATCAGCAAAAGCTTACATCGGCAAGTGAACAAAGCTCC 1559
 Db 26894 TCTCAAGCGAATAATAGGAGATCAGCAAAAGCTTACATCGGCAAGTGAACAAAGCTCC 26835
 QY 1560 ACAGTTGGCTATCCTTATGTCGAATTTCTAATGACTCTACACAGAGTTGATAGCTAA 1619
 Db 26834 ACAGTTGGCTATCCTTATGTCGAATTTCTAATGACTCTACACAGAGTTGATAGCTAA 26775
 QY 1620 ATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAAGCACTTTCCCT 1679
 Db 26774 ATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAAGCACTTTCCCT 26715
 QY 1680 TGAACGAACTCCTTGTATTAATCAGAGGTGCTGCTCAATATCGGCTCTCTATAATCTGG 1739
 Db 26714 TGAACGAACTCCTTGTATTAATCAGAGGTGCTGCTCAATATCGGCTCTCTATAATCTGG 26655
 QY 1740 TTATCTCCAATAACACACCTAAGTGTTCGTTTGGAGAGATTTATGTCGTTTGGTAAG 1799
 Db 26654 TTATCTCCAATAACACACCTAAGTGTTCGTTTGGAGAGATTTATGTCGTTTGGTAAG 26595
 QY 1800 GCCTTTGTGAGGCTTTACCAACACACTAGAACGATCTTCAATAATAAAGA 1852
 Db 26594 GCCTTTGTGAGGCTTTACCAACACACTAGAACGATCTTCAATAATAAAGA 26542

RESULT 4

AA557031
 ID AA557031 standard; DNA; 1956 BP.
 XX
 AC AA557031;
 XX
 DT 11-SEP-2003 (revised)
 DT 16-JAN-2002 (first entry)
 XX
 DE C. pneumoniae DNA encoding the CT622 homologue CPn0728.
 XX
 KW Chlamydia; ds; sexually transmitted disease; PID; antibacterial;
 KW pelvic inflammatory disease; antigen; trachoma; gynecological;
 KW acute respiratory tract infection; atherosclerosis; male infertility;
 KW coronary heart disease

Chlamydia pneumoniae.

WO2000181379-A2.

01-NOV-2001.

23-APR-2001; 2001WO-US013081.

21-APR-2000; 2000US-0198853P.

20-JUL-2000; 2000US-0219752P.

(CORI), CORIXA CORP.

Bhatia A, Probst-P, Stromberg EJ;

WPI; 2001-616771/71.

P-PSDB; AAU38899.

PT New polynucleotide for treating Chlamydia infections encodes a
 XX polynucleotides containing an immunogenic portion of a Chlamydia antigen.
 PS Disclosure; Page 143; 208pp; English.

CC The invention relates to isolated polynucleotide encoding at least a
 CC partial Chlamydia protein which is an antigenic fragment, or the
 CC complements, fragments, homologues and variants, and antibodies raised
 CC against the antigenic proteins (or fragments). The nucleic acids,
 CC proteins and antibodies are used to diagnose and treat Chlamydia
 CC infections (e.g. a sexually transmitted disease, pelvic inflammatory
 CC disease (PID), acute respiratory tract infection, trachoma,
 CC atherosclerosis and coronary heart disease) in a patient, and in the
 CC treatment of male infertility. The compounds of the invention are also
 CC useful for detecting the presence of Chlamydia in a patient, and
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC The present sequence encodes a Chlamydia antigen. (Updated on 11-SEP-2003
 CC to standardise OS field)

XX
 SQ Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;

Query Match 94.4%; Score 1748.8; DB 4; Length 1956;

Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATTCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
 Db 205 ATGAGTCTGGCAGATAAGCTGGGTATTCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264
 QY 61 TCTGCAGAGCTGGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCACGTTTGTAT 120
 Db 265 TCTGCAGAGCTGGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCACGTTTGTAT 324
 QY 121 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTTTCTCTCAACATCATAGCT 180
 Db 325 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTTTCTCTCAACATCATAGCT 384
 QY 181 GACATACAGGCTGCTTTGGTGAGGCTCCAGATGCTGTCATTAATAAAGGATACACG 240
 Db 385 GACATACAGGCTGCTTTGGTGAGGCTCCAGATGCTGTCATTAATAAAGGATACACG 444
 QY 241 GCTACTGATGAGGAACCGCAATCGCTCGGAGTGGGAACTAAGATGCCGATGCAGTT 300
 Db 445 GCTACTGATGAGGAACCGCAATCGCTCGGAGTGGGAACTAAGATGCCGATGCAGTT 504
 QY 301 AAAGTTGGCGGCAAAATTACAGAAATTAGCGAAATATGCTCGGATTAACCAAGCGATTCTT 360
 Db 505 AAAGTTGGCGGCAAAATTACAGAAATTAGCGAAATATGCTCGGATTAACCAAGCGATTCTT 564
 QY 361 GACTCTTTAGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 420
 Db 565 GACTCTTTAGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 624
 QY 421 GCAACCAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGTCCCA 480
 Db 625 GCAACCAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGTCCCA 684
 QY 481 GGGAAAAACGCTGCAATTGCTTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
 Db 685 GGGAAAAACGCTGCAATTGCTTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 744
 QY 541 ATAGAGAAAGATGGAATGCGATTAGGGATGCATATTTTTCAGACAGACGCTAGTGGGA 600
 Db 745 ATAGAGAAAGATGGAATGCGATTAGGGATGCATATTTTTCAGACAGACGCTAGTGGGA 804
 QY 601 GCTGTAGAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTTCAGCTTAAGCAGCA 660
 Db 805 GCTGTAGAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTTCAGCTTAAGCAGCA 864
 QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTTCCCGGACTCTTCCAATT 720
 Db 865 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTTCCCGGACTCTTCCAATT 924
 QY 721 CTTCAAGAGCGGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 780
 Db 925 CTTCAAGAGCGGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 984
 QY 781 GCAGATGGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGACACACAGGA 840

Db 985 GCAGATGGTTCGTGATGTTCCAAATCCAGAACTACAGATTGGAGCTCCAAAGCAACAGGA 1044
QY 841 AGTAGATTGGTAGTATTCGTGTTTCCATGCTGTAGATGATGCTGAAAATGAGCCGCT 900
Db 1045 AGTAGATTGGTAGTATTCGTGTTTCCATGCTGTAGATGATGCTGAAAATGAGCCGCT 1104
QY 901 TCCATTTTGTATGCTGGGTTTCCTCAGATGATTCACATGTTCAATACGGAATAATCCTGAT 960
Db 1105 TCCATTTTGTATGCTGGGTTTCCTCAGATGATTCACATGTTCAATACGGAATAATCCTGAT 1164
QY 961 TCTCAAGCTCCCAACAGGAGCTCGCAGCAGCTAGACGAGCGAAAGCGCTGGAGAT 1020
Db 1165 TCTCAAGCTCCCAACAGGAGCTCGCAGCAGCTAGACGAGCGAAAGCGCTGGAGAT 1224
QY 1021 GACAGTCTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1080
Db 1225 GACAGTCTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1284
QY 1081 GCTGGGCAACACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG 1140
Db 1285 GCTGGGCAACACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG 1344
QY 1141 AGCGCAGAGTCTCTCCGCTCGCAGCAAGTCTATAGGTCATCTGTAABACAGCTTTAC 1200
Db 1345 AGCGCAGAGTCTCTCCGCTCGCAGCAAGTCTATAGGTCATCTGTAABACAGCTTTAC 1404
QY 1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGCT 1260
Db 1405 AAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGCT 1464
QY 1261 TACAAATCCATCATGATGCTATGTTAGGACGCAAAATGATGCGACTCGTGATGTGATA 1320
Db 1465 TACAAATCCATCATGATGCTATGTTAGGACGCAAAATGATGCGACTCGTGATGTGATA 1524
QY 1321 AACAAATGATAGTACCCCGCTCTCACAGATCCCTTCTAGAGCAGCAACAGAAAGCTCGA 1380
Db 1525 AACAAATGATAGTACCCCGCTCTCACAGATCCCTTCTAGAGCAGCAACAGAAAGCTCGA 1584
QY 1381 GGACCCAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGAACTCTTT 1440
Db 1585 GGACCCAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGAACTCTTT 1644
QY 1441 GGAGATGCTATAGTCAAGTTTCGGACTACAAATCTGTATGAGATCACTCAGTCAAT 1500
Db 1645 GGAGATGCTATAGTCAAGTTTCGGACTACAAATCTGTATGAGATCACTCAGTCAAT 1704
QY 1501 CCTCAAGCGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCCCTCCA 1560
Db 1705 CCTCAAGCGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCCCTCCA 1764
QY 1561 CAGTTTGGCTATCCTTATGTGCAACTTTCTTAATGACTCTACACAGAAAGTTTCATAGTAAA 1620
Db 1765 CAGTTTGGCTATCCTTATGTGCAACTTTCTTAATGACTCTACACAGAAAGTTTCATAGTAAA 1824
QY 1621 TTGAAAGTTTGTGTTGGTGGAGGATCTAGGACAGCTGAAATAAAGCACTTTCCTTT 1680
Db 1825 TTGAAAGTTTGTGTTGGTGGAGGATCTAGGACAGCTGAAATAAAGCACTTTCCTTT 1884
QY 1681 GAAACGAACTCCTTGTGTTTATTCAGCAGGTCGCTCAATATCGGCTCTCTATATCTGTT 1740
Db 1885 GAAACGAACTCCTTGTGTTTATTCAGCAGGTCGCTCAATATCGGCTCTCTATATCTGTT 1944
QY 1741 TATCTCCAAATA 1752
Db 1945 TATCTCCAAATA 1956

RESULT 5
ABL92612
ID ABL92612 standard; DNA; 1956 BP.
XX
AC ABL92612;

XX 29-AUG-2003 (revised)
DT 05-JUN-2002 (first entry)
XX Chlamydia pneumoniae DNA sequence SEQ ID NO:385.
XX Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
XX antibacterial; immunostimulant; immune response;
XX Chlamydia-specific T-cell response; gene; ds.
XX Chlamydia-specific T-cell response; gene; ds.
OS Chlamydia pneumoniae.
XX WC200208267-A2.
XX 31-JAN-2002.
XX 20-JUL-2001; 2001WO-US023121.
XX 20-JUL-2000; 2000US-00620412.
XX 23-APR-2001; 2001US-00841132.
XX (CORI-) CORIXA CORP.
XX Fling SP, Skeiky YAW, Probst P, Bhatia A;
XX WPI; 2002-179901/23.
XX Novel compositions comprising Chlamydia Cap1 protein and its use in the
XX treatment of Chlamydia infection.
XX Disclosure; Page 336-337; 537pp; English.
XX The present invention describes compositions comprising a Chlamydia Cap1
XX protein and methods for the diagnosis and therapy of Chlamydia infection.
XX Chlamydia DNA and protein sequences from the present invention can have
XX antibacterial and immunostimulant activities, and can be used in
XX vaccines. Compounds from the present invention can be used for eliciting
XX an immune response, specifically stimulating a Chlamydia-specific T-cell
XX response or inhibiting the development of a Chlamydia infection in an
XX animal. Methods from the present invention can be used for detecting the
XX presence of Chlamydia in a patient; to stimulate and/or expand T cells
XX specific for a Chlamydia protein; and for treatment of a Chlamydia
XX infection. ABL92394 to ABL92709 and ABL94096 to ABL94374 represent
XX sequences used in the exemplification of the present invention. (Updated
XX on 29-AUG-2003 to standardise OS field)
SQ Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;
Query Match 94.4%; Score 1748.8; DB 6; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGCTCTTCTACTAGCAGA 60
Db 205 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGCTCTTCTACTAGCAGA 264
QY 61 TCTGCGACGTGAGTCAACGACGACGCGACCTACGCTCCTCCACCCAGCTTTGAT 120
Db 265 TCTGCGACGTGAGTCAACGACGACGCGACCTACGCTCCTCCACCCAGCTTTGAT 324
QY 121 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCAGTACT 180
Db 325 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCAGTACT 384
QY 181 GACATACAGCTCTTTTGGTGAGCCCTCCAGGATGCTGTCACTAATAAAGGATACAGCG 240
Db 385 GACATACAGCTCTTTTGGTGAGCCCTCCAGGATGCTGTCACTAATAAAGGATACAGCG 444
QY 241 GCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCGATGAGTT 300
Db 445 GCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCGATGAGTT 504
QY 301 AAAGTTGGCGCGCAATTACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTTCT 360

PT Novel Chlamydia pneumoniae protein useful in the manufacture of a
 PT medicament for treatment or prevention of infection due to Chlamydia,
 XX preferably Chlamydia pneumoniae, and for diagnostic purposes.
 XX

PS Claim 5; Page 47-48; 364pp; English.

XX Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
 CC pneumoniae (strain CW029), and ABL91184-ABL91373 represent DNA encoding
 CC them. The proteins are predicted to be immunogenic and may therefore be
 CC useful in vaccine production and for diagnostic purposes. Chlamydia
 CC pneumoniae is a common cause of respiratory disease in humans, and is
 CC also involved in the development of cardiovascular diseases such as
 CC atherosclerosis, coronary artery disease, carotid artery stenosis,
 CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
 CC claudication and stroke. The proteins and nucleic acids of the invention
 CC may be used in vaccines and pharmaceutical compositions for the
 CC prevention or treatment of chlamydial infections, particularly Chlamydia
 CC pneumoniae infections. The proteins may also be used in the detection of
 CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
 CC DNA probe assay or blotting techniques for determining Chlamydia
 CC pneumoniae gene expression. The present sequence represents a
 CC specifically claimed DNA which encodes a Chlamydia pneumoniae protein of
 CC the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;

Query Match 34.4%; Score 1748.8; DB 6; Length 1956;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGCGAGATAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTACTACGAGA 60
 DB 205 ATGAGTCTGCGAGATAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTACTACGAGA 264

QY 61 TCTGCGACGTGGACTCAACGACAGCAGCAGCCGACCTACGCGCTCCTCCACCCAGCTTTGAT 120
 DB 265 TCTGCGACGTGGACTCAACGACAGCAGCAGCCGACCTACGCGCTCCTCCACCCAGCTTTGAT 324

QY 121 GATTATAGACTCAGCGGCAACAGCTTAGATACTATCTTTACCTCAACATCACTAGCT 180
 DB 325 GATTATAGACTCAGCGGCAACAGCTTAGATACTATCTTTACCTCAACATCACTAGCT 384

QY 181 GACATACAGGCTGCTTGGTGGAGCTCCAGGATGCTCTACTAATAATAAAGGATACAGCG 240
 DB 385 GACATACAGGCTGCTTGGTGGAGCTCCAGGATGCTCTACTAATAATAAAGGATACAGCG 444

QY 241 GCTACTGATGAGGAACCGCAATGCTGCGGAGTGGGAAACTAAGATGCGGATGCGATT 300
 DB 445 GCTACTGATGAGGAACCGCAATGCTGCGGAGTGGGAAACTAAGATGCGGATGCGATT 504

QY 301 AAGTTGGCGCGCAANTTACAGATTTAGCGAATATGCTTCGGNTAACCAAGCGATTCTT 360
 DB 505 AAGTTGGCGCGCAANTTACAGATTTAGCGAATATGCTTCGGNTAACCAAGCGATTCTT 564

QY 361 GACTCTTTAGGTAACAGTACTCTCTTGACCTCTTACAGGCTGCTCTTCTCCCAATCTGTA 420
 DB 565 GACTCTTTAGGTAACAGTACTCTCTTGACCTCTTACAGGCTGCTCTTCTCCCAATCTGTA 624

QY 421 GCAACATTAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
 DB 625 GCAACATTAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 684

QY 481 GGGAAACCGCTGCAATGCTCAATCTTTAGTTGANTCAGACAGNTGCTACAGCGACACAG 540
 DB 685 GGGAAACCGCTGCAATGCTCAATCTTTAGTTGANTCAGACAGNTGCTACAGCGACACAG 744

QY 541 ATAGAGAAAGATGGAATGCGATTAGGATGCAATTTTTCAGGACAGAACCGCTAGTGGG 600
 DB 745 ATAGAGAAAGATGGAATGCGATTAGGATGCAATTTTTCAGGACAGAACCGCTAGTGGG 804

QY 601 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGAGTAAAGCAGCA 660
 DB 805 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGAGTAAAGCAGCA 864

QY 661 ATCGTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAATT 720
 DB 865 ATCGTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAATT 924

QY 721 CTTCAAGAGCGGAACAAATGGTAATAACAGGCTGAGAAAAGTCTTAAAATATCAAACCT 780
 DB 925 CTTCAAGAGCGGAACAAATGGTAATAACAGGCTGAGAAAAGTCTTAAAATATCAAACCT 984

QY 781 GCAGATGGTCTGATGCTTCCAAATCCAGGAACCTACAGTTTGGAGGCTCCAAAGCAACAGGA 840
 DB 985 GCAGATGGTCTGATGCTTCCAAATCCAGGAACCTACAGTTTGGAGGCTCCAAAGCAACAGGA 1044

QY 841 AGTAGTATGGTAGTATTCGTGTTCCATGCTGTTTAGATGATGCTGAAAATAGACCGCT 900
 DB 1045 AGTAGTATGGTAGTATTCGTGTTCCATGCTGTTTAGATGATGCTGAAAATAGACCGCT 1104

QY 901 TCCATTTTGTATGCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATCTCTGAT 960
 DB 1105 TCCATTTTGTATGCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATCTCTGAT 1164

QY 961 TCTCAAGCTGCCCAACAGGAGCTCGCAGCACAAAGCTTAGAGCAGCGAAAGCCGCTGAGAT 1020
 DB 1165 TCTCAAGCTGCCCAACAGGAGCTCGCAGCACAAAGCTTAGAGCAGCGAAAGCCGCTGAGAT 1224

QY 1021 GACAGTGTGCTGACGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGCGCTCTAGGTAA 1080
 DB 1225 GACAGTGTGCTGACGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGCGCTCTAGGTAA 1284

QY 1081 GCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTCTCTGCTGTGTG 1140
 DB 1285 GCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTCTCTGCTGTGTG 1344

QY 1141 AGCGCAGGAGTTCTCTCCGCTGCGCAGCAAGTCTTAGGGTCTATCTGTAAAACAGCTTTAC 1200
 DB 1345 AGCGCAGGAGTTCTCTCCGCTGCGCAGCAAGTCTTAGGGTCTATCTGTAAAACAGCTTTAC 1404

QY 1201 AAGACCTCAAAATCTACAGGTTCTGATATAAAACACAGATATCAGCAGGTTATGATGT 1260
 DB 1405 AAGACCTCAAAATCTACAGGTTCTGATATAAAACACAGATATCAGCAGGTTATGATGT 1464

QY 1261 TACAAATCCATCAATGATGCTTATGGTGGCAGCAAGTATGATGCGACTCGTGTGTGATA 1320
 DB 1465 TACAAATCCATCAATGATGCTTATGGTGGCAGCAAGTATGATGCGACTCGTGTGTGATA 1524

QY 1321 AACAAATGTAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA 1380
 DB 1525 AACAAATGTAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA 1584

QY 1381 GGACCAAGAAAACAGATCAAGCCCTCGTAGGGTGAATTTCTGGCAATAGCAGAACTCTT 1440
 DB 1585 GGACCAAGAAAACAGATCAAGCCCTCGTAGGGTGAATTTCTGGCAATAGCAGAACTCTT 1644

QY 1441 GGAGATGCTCTAGTCAAGTTTCGGCCTACCAATCTGTAATGCAATCACTCAGTCCGAAT 1500
 DB 1645 GGAGATGCTCTAGTCAAGTTTCGGCCTACCAATCTGTAATGCAATCACTCAGTCCGAAT 1704

QY 1501 CTTCAAGCGAATAAGAGAGATCAGACAAAAGCTTTACATCGGCAAGTGAACAAGCTCCCA 1560
 DB 1705 CTTCAAGCGAATAAGAGAGATCAGACAAAAGCTTTACATCGGCAAGTGAACAAGCTCCCA 1764

QY 1561 CAGTTTGGGTATCTTATGTCGAACTTTCTATAGTCTTACACAGAGTTCTATAGTAAA 1620
 DB 1765 CAGTTTGGGTATCTTATGTCGAACTTTCTATAGTCTTACACAGAGTTCTATAGTAAA 1824

QY 1621 TTAGAAAGTTTCTTGTGCTGAAGGATCTAGGACAGCAGCTGTAATAAAGCACTTTTCCCTTT 1680
 DB 1825 TTAGAAAGTTTCTTGTGCTGAAGGATCTAGGACAGCAGCTGTAATAAAGCACTTTTCCCTTT 1884

QY 1681 GAAACGAACTCTCTGTTTATTCAGCAGGTCGTGGTCAATATATCGGCTCTCTATATTTCTGTT 1740
 DB 1885 GAAACGAACTCTCTGTTTATTCAGCAGGTCGTGGTCAATATATCGGCTCTCTATATTTCTGTT 1944

QY 1741 TATCTCCATATA 1752
 Db 1945 TATCTCCATATA 1956
 RESULT 7
 ID ADD42825 standard; DNA; 1956 BP.
 XX AC ADD42825;
 XX DT 15-JAN-2004 (first entry)
 XX Chlamydia pneumoniae antigen polynucleotide SEQ ID NO:63.
 DE Chlamydia pneumoniae; Chlamydia; antibiotic; antiinflammatory;
 KW antinfertility; cardant; antiarteriosclerotic; ophthalmological;
 KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
 KW tubal obstruction; infertility; male infertility; ocular infection;
 KW blindness; acute respiratory tract infection; atherosclerosis;
 KW coronary heart disease; gene; ds.
 XX Chlamydia pneumoniae.
 OS WO2003041560-A2.
 XX PN 22-MAY-2003.
 XX PD 05-NOV-2002; 2002WO-US035624.
 XX PF 06-NOV-2001; 2001US-00012256.
 XX PR 05-DEC-2001; 2001US-00007693.
 XX PR 15-JUL-2002; 2002US-00197220.
 XX PA (CORI-) CORIXA CORP.
 XX Bhatia A, Guiderian J, Skeiky YAW, Maisonneuve JL, Barth B;
 PI Probst P;
 XX WPI: 2003-441771/41.
 DR P-PSDB; ADD42865.
 XX New DNA and proteins comprising a portion of a Chlamydia antigen, useful
 PT for diagnosing or treating Chlamydial infections, particularly as
 PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic
 PT inflammatory disease.
 XX Disclosure; SEQ ID NO 63; 275pp; English.
 XX The present invention describes compounds and methods for diagnosing and
 CC treating Chlamydial infection. Chlamydia polynucleotide and protein
 CC sequences have antibiotic, antiinflammatory, antiinfertility, cardiac,
 CC antiarteriosclerotic and ophthalmological activities, and can be used in
 CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
 CC compositions or methods from the present invention can be used for the
 CC serodiagnosis or treatment of Chlamydial infections, particularly in
 CC humans. The polynucleotides, proteins or compositions are particularly
 CC useful for stimulating an immune response in a patient, or for
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC Specifically, the polynucleotides, proteins or compositions are useful as
 CC vaccines for treating or preventing Chlamydial infections including
 CC pelvic inflammatory disease (which results in tubal obstruction and
 CC infertility in women), male infertility, ocular infection (which may
 CC cause blindness), acute respiratory tract infections, atherosclerosis, or
 CC coronary heart disease. The present sequence is used in the
 CC exemplification of the present invention.
 XX Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;
 Query Match 94.4%; Score 1748.8; DB 9; Length 1956;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 60
 Db 205 ATGAGCTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 264
 QY 61 TCTGCAGACGTGGACTCAACGACAGCGACCGCACCTACGCTCCCTCCACCCACGTTTGTAT 120
 Db 265 TCTGCAGACGTGGACTCAACGACAGCGACCGCACCTACGCTCCCTCCACCCACGTTTGTAT 324
 QY 121 GATTATAAGACTCAAGCGCAAAACAGCTTACGATACCTATCTTTACCTCAACATCACTAGCT 180
 Db 325 GATTATAAGACTCAAGCGCAAAACAGCTTACGATACCTATCTTTACCTCAACATCACTAGCT 384
 QY 181 GACATACAGGCTGCTTTGGTGGAGCTCCAGAGTGTCTGCTCACTAATATAAGGATACAGCG 240
 Db 385 GACATACAGGCTGCTTTGGTGGAGCTCCAGAGTGTCTGCTCACTAATATAAGGATACAGCG 444
 QY 241 GCTACTGATAGGAACCGCAATCGCTGGAGTGGGAACTAAGAACTGCGGATGCGGATGCGG 300
 Db 445 GCTACTGATAGGAACCGCAATCGCTGGAGTGGGAACTAAGAACTGCGGATGCGGATGCGG 504
 QY 301 AAAGTTGGCGGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 360
 Db 505 AAAGTTGGCGGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 564
 QY 361 GACTCTTTAGGTAAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAACTCTGTA 420
 Db 565 GACTCTTTAGGTAAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAACTCTGTA 624
 QY 421 GCATAAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCCAGTAGTCCCA 480
 Db 625 GCATAAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCCAGTAGTCCCA 684
 QY 481 GGGAAACGCTGCAATTTGCTCAATCTTTAGTTGATGACAGAGATGCTGACGCGACAG 540
 Db 685 GGGAAACGCTGCAATTTGCTCAATCTTTAGTTGATGACAGAGATGCTGACGCGACAG 744
 QY 541 ATAGAGAAAGATGGAATGCGATAGGATGCGATATTTTTCAGGACGACGAGCTAGTCCA 600
 Db 745 ATAGAGAAAGATGGAATGCGATAGGATGCGATATTTTTCAGGACGACGAGCTAGTCCA 804
 QY 601 GCTGTAGAAAATGCTTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTTAAGACGCA 660
 Db 805 GCTGTAGAAAATGCTTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTTAAGACGCA 864
 QY 661 ATCGTACTGCTAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAAAT 720
 Db 865 ATCGTACTGCTAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAAAT 924
 QY 721 CTTCAAGAGCGGAAACAAATGCTTAATACAGGCTCAGAAAGATCTTAAAAATCAAACT 780
 Db 925 CTTCAAGAGCGGAAACAAATGCTTAATACAGGCTCAGAAAGATCTTAAAAATCAAACT 984
 QY 781 GCAGATGGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGGA 840
 Db 985 GCAGATGGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGGA 1044
 QY 841 AGTAGTATTGTTAGTATTTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT 900
 Db 1045 AGTAGTATTGTTAGTATTTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT 1104
 QY 901 TCCATTTTGTATGCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATTCCTGAT 960
 Db 1105 TCCATTTTGTATGCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATTCCTGAT 1164
 QY 961 TCTCAAGTGTCCCAACAGGAGCTCGCAGCAACAGCTAGAGCAGGAAAGCGCTGGAGAT 1020
 Db 1165 TCTCAAGTGTCCCAACAGGAGCTCGCAGCAACAGCTAGAGCAGGAAAGCGCTGGAGAT 1224
 QY 1021 GACAGTGTCTGCGAGCGCTGGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1080
 Db 1225 GACAGTGTCTGCGAGCGCTGGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1284
 QY 1081 GCTGGCAACACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTG 1140

Db 1285 GCTGGCAACACAGGGGATCTCAATGCTTTAGACAGATCGCTTCGCTGCTGTG 1344
Qy 1141 AGCGCAGAGTTCCTCCGCTGCGACGAGTCTATAGGCTCATCTGTAACACAGCTTTAC 1200
Db 1345 AGCGCAGAGTTCCTCCGCTGCGACGAGTCTATAGGCTCATCTGTAACACAGCTTTAC 1404
Qy 1201 AAGACCTCAAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
Db 1405 AAGACCTCAAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATGATGCT 1464
Qy 1261 TACAATCCATCAATGATGCTATGCTAGGCGACGAAATGATGCGACTCGTGATGATGATA 1320
Db 1465 TACAATCCATCAATGATGCTATGCTAGGCGACGAAATGATGCGACTCGTGATGATGATA 1524
Qy 1321 AACATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGACGACGAAACAGAGCTCGA 1380
Db 1525 AACATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGACGACGAAACAGAGCTCGA 1584
Qy 1381 GGACGAGAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGACTCTT 1440
Db 1585 GGACGAGAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGACTCTT 1644
Qy 1441 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCAGATCACTCAGTCCGAAT 1500
Db 1645 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCAGATCACTCAGTCCGAAT 1704
Qy 1501 CCTCAAGGCAATATAGAGAGATCAGACAAAGCTTACATCGGAGTACAAAGCTTCCA 1560
Db 1705 CCTCAAGGCAATATAGAGAGATCAGACAAAGCTTACATCGGAGTACAAAGCTTCCA 1764
Qy 1561 CAGTTGGCTATCCTTATGTCACCTTCTATGACTCTACACAGAGTTCATAGCTAAA 1620
Db 1765 CAGTTGGCTATCCTTATGTCACCTTCTATGACTCTACACAGAGTTCATAGCTAAA 1824
Qy 1621 TTAGAAAGTTTGTTCGCTGAAGGATCTAGGACAGCAGCTGAAATATAAGCACTTTCCTTT 1680
Db 1825 TTAGAAAGTTTGTTCGCTGAAGGATCTAGGACAGCAGCTGAAATATAAGCACTTTCCTTT 1884
Qy 1681 GAACGAACTCTGTTTATTCAGCAGAGTGTGCTCAATTCGGCTCTCTATATCTGGT 1740
Db 1885 GAACGAACTCTGTTTATTCAGCAGAGTGTGCTCAATTCGGCTCTCTATATCTGGT 1944
Qy 1741 TATCTCCAAATA 1752
Db 1945 TATCTCCAAATA 1956

RESULT 8

AD02066

ID AD02066 standard; DNA; 2238 BP.

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

XX AC AD02066;

FT misc_feature /product= "Truncated Chlamydia pneumoniae 76KDa protein"
FT FT 2122..2238
FT FT /*tag= C
FT FT /note= "This part of the sequence is unrelated to C.
XX XX pneumoniae 76 kDa gene"
PN W0200066739-A2.
XX 09-NOV-2000.
XX 03-MAY-2000; 2000WO-CA000511.
XX 03-MAY-1999; 99US-0132270P.
XX 30-JUN-1999; 99US-0141276P.
XX (AVET) AVENTIS PASTEUR LTD.
XX Murdin AD, Comen RP, Wang J, Dunn P;
XX WPI; 2000-687542/67.
XX P-PSDB; AAY11957.
XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
XX for vaccinating against Chlamydia infections.
XX Claim 32; Fig 3; 112pp; English.
XX The present sequence is a DNA coding for a fusion protein comprising a
XX truncated Chlamydia pneumoniae 76 kDa protein and vector-encoded
XX residues. C. pneumoniae 76 kDa protein is used in the diagnosis,
XX prevention and treatment of C. pneumoniae infections (e.g. pneumonia,
XX upper respiratory tract disease, bronchitis, sinusitis and acute
XX respiratory disease such as cough, sore throat, hoarseness, fever; and
XX abnormal chest sounds on auscultation). C. pneumoniae sequence is also
XX used as vaccines for immunising humans against diseases caused by C.
XX pneumoniae
SQ Sequence 2238 BP; 689 A; 459 C; 497 G; 593 T; 0 U; 0 Other;

Query Match 62.7%; Score 1161.8; DB 3; Length 2238;
Best Local Similarity 99.7%; Pred. No. 8.3e-314;
Matches 1174; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 ATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 60
Db 970 ATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 1029
Qy 61 TCTGCAGACGTGGACTCAACGACGACGACGACGACGACGACGACGACGACGACGACGACG 120
Db 1030 TCTGCAGACGTGGACTCAACGACGACGACGACGACGACGACGACGACGACGACGACG 1089
Qy 121 GATTATAAGACTCAAGCGGAAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT 180
Db 1090 GATTATAAGACTCAAGCGGAAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT 1149
Qy 181 GACATACAGCTGCTTTGGTGAGCTCCAGATGCTCTCACTAATAATAAGGATACAGCG 240
Db 1150 GACATACAGCTGCTTTGGTGAGCTCCAGATGCTCTCACTAATAATAAGGATACAGCG 1209
Qy 241 GCTACTGATGAGGAAACCGCAATCGCTGCGAGTGGGAAACCTAAGATGCCGATGAGTT 300
Db 1210 GCTACTGATGAGGAAACCGCAATCGCTGCGAGTGGGAAACCTAAGATGCCGATGAGTT 1269
Qy 301 AAGATTGGCGGCAATATCAGATTTAGCGAATATGCTTCGGATACCAAGGATCTTT 360
Db 1270 AAGATTGGCGGCAATATCAGATTTAGCGAATATGCTTCGGATACCAAGGATCTTT 1329
Qy 361 GACTCTTTTAGTAACTGACTTCTCTCGACTCTTTACAGCTGCTCTTCTCCCAATCTGTA 420
Db 1330 GACTCTTTTAGTAACTGACTTCTCTCGACTCTTTACAGCTGCTCTTCTCCCAATCTGTA 1389
Qy 421 GCAACAAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATACCAAGTGTCCCA 480

Db 1390 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA 1449
Qy 481 GGGAAAAGCCCTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCAGACAG 540
Db 1450 GGGAAAAGCCCTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCAGACAG 1509
Qy 541 ATAGAGAAAGATGGAAATGCGATTAGGGATGCATATTTGACAGACAGAACCGCTAGTGG 600
Db 1510 ATAGAGAAAGATGGAAATGCGATTAGGGATGCATATTTGACAGACAGAACCGCTAGTGG 1569
Qy 601 GCTGTAGAAAATGCTAAATCTTAATACAGTATTAAGCAACATAGATTCAGCTAAAGCAGCA 660
Db 1570 GCTGTAGAAAATGCTAAATCTTAATACAGTATTAAGCAACATAGATTCAGCTAAAGCAGCA 1629
Qy 661 ATCGTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCTCAATT 720
Db 1630 ATCGTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCTCAATT 1689
Qy 721 CTTCAAGAAAGCGGAAACAAATGCTAATACAGGCTGAGAAAGATCTTTAAAAATATCAAACT 780
Db 1690 CTTCAAGAAAGCGGAAACAAATGCTAATACAGGCTGAGAAAGATCTTTAAAAATATCAAACT 1749
Qy 781 GCAGATGCTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAGCAACAAGGA 840
Db 1750 GCAGATGCTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAGCAACAAGGA 1809
Qy 841 AGTAGTATGCTAGTATGCTGTTTCCATGCTGTTAGATGATGCTGAAATAGAGCCCT 900
Db 1810 AGTAGTATGCTAGTATGCTGTTTCCATGCTGTTAGATGATGCTGAAATAGAGCCCT 1869
Qy 901 TCCATTTTGTGCTGCTGGTTCGTCAGATGATTCACATGTTCAATACGAAAATCCTGAT 960
Db 1870 TCCATTTTGTGCTGCTGGTTCGTCAGATGATTCACATGTTCAATACGAAAATCCTGAT 1929
Qy 961 TCTCAAGCTGCCAACAGAGAGCTGCAGACACAAGCTAGACAGCAAGCCGCTGGAGAT 1020
Db 1930 TCTCAAGCTGCCAACAGAGAGCTGCAGACACAAGCTAGACAGCAAGCCGCTGGAGAT 1989
Qy 1021 GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTTAGGTAAA 1080
Db 1990 GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTTAGGTAAA 2049
Qy 1081 GCTGGGCAACCAACAGGGGATCTCAATGCTTTAGACAGATGCTGCTGCTGCTGCTG 1140
Db 2050 GCTGGGCAACCAACAGGGGATCTCAATGCTTTAGACAGATGCTGCTGCTGCTGCTG 2109
Qy 1141 AGCGCAGAGTTCCTCCCGCTGCGAGCAAGTTCATAG 1177
Db 2110 AGCGCAGAGTTCCTCCCGCTGCGAGCAAGTTCATAG 2145

RESULT 9

AA02065
ID AA02065 standard; DNA; 1456 BP.
XX AC AA02065;
XX DT 26-MAR-2001 (first entry)
XX DE 3'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.
XX KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
KW upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; treatment; truncation mutant; ds.
XX Chlamydia pneumoniae.
OS Synthetic.
XX Key Location/Qualifiers
FH 101..1456
CDS /*tag= a
FT /product= "3'-truncated Chlamydia pneumoniae 76kDa
FT

protein"
/note= "The coding region does not include stop codon"
/partial

W0200066739-A2.

XX 09-NOV-2000.

XX 03-MAY-2000; 2000WO-CA000511.

XX 03-MAY-1999; 99US-0132270P.

XX 30-JUN-1999; 99US-0141276P.

XX (AVET) AVENTIS PASTEUR LTD.

XX Murdin AD, Oomen RP, Wang J, Dunn P;

XX WPI; 2000-687542/67.

XX P-PSDB; AAY71956.

Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
for vaccinating against Chlamydia infections.

XX Claim 2c; Page 102-104; 11:2pp; English.

The present sequence is a DNA coding for 3'-truncated Chlamydia
pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the
diagnosis, prevention and treatment of C. pneumoniae infections (e.g.
pneumonia, upper respiratory tract disease, bronchitis, sinusitis and
acute respiratory disease such as cough, sore throat, hoarseness, fever;
and abnormal chest sounds on auscultation). C. pneumoniae sequence is
also used as vaccines for immunising humans against diseases caused by C.
pneumoniae

XX Sequence 1456 BP; 452 A; 308 C; 331 G; 365 T; 0 U; 0 Other;

Query Match 62.1%; Score 1151; DB 3; Length 1456;

Best Local Similarity 100.0%; Pred. No. 6.6e-311;

Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 60
Db 305 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 364
Qy 61 TCTGCAGACGTGGACTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120
Db 365 TCTGCAGACGTGGACTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 424
Qy 121 GATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACTAGCT 180
Db 425 GATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACTAGCT 484
Qy 181 GACATACAGCTGCTTTGGTGAGCCTCCAGAGTCTGTCACTAATAAAGGATACAGCG 240
Db 485 GACATACAGCTGCTTTGGTGAGCCTCCAGAGTCTGTCACTAATAAAGGATACAGCG 544
Qy 241 GCTACTGATGAGGAACCGCAATCGCTGCGAGTGGGAAACTAAGAAATGCGGATG 300
Db 545 GCTACTGATGAGGAACCGCAATCGCTGCGAGTGGGAAACTAAGAAATGCGGATG 604
Qy 301 AAAGTTGGCGGCAATTACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATTCCT 360
Db 605 AAAGTTGGCGGCAATTACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATTCCT 664
Qy 361 GACTCTTTAGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCATCTGTA 420
Db 665 GACTCTTTAGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCATCTGTA 724
Qy 421 GCATAAATTAACAAAGCAGCTGAGCTTCTTAAAGAGATGAAGATAACCCAGTAGTCCCA 480
Db 725 GCATAAATTAACAAAGCAGCTGAGCTTCTTAAAGAGATGAAGATAACCCAGTAGTCCCA 784
Qy 481 GGGAAAACGCTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCAGACAG 540

Db	785	GGGAAAAGCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACGCGACACAG	844
Qy	541	ATAGAGAAAGATGGAAATGCGATTAGGGATGCATATTTTGCAGGACAGAACGCTAGTGG	600
Db	845	ATAGAGAAAGATGGAAATGCGATTAGGGATGCATATTTTGCAGGACAGAACGCTAGTGG	904
Qy	601	GCTCTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGAGCTAAAGCAGCA	660
Db	905	GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGAGCTAAAGCAGCA	964
Qy	661	ATCGCTACTGCTAAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACCTCTCCAATT	720
Db	965	ATCGCTACTGCTAAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACCTCTCCAATT	1024
Qy	721	CTTCAAGAACGGGAACAAATGTTATACAGGCTGAGAAGAGTCTTAAAAAATATCAAACT	780
Db	1025	CTTCAAGAACGGGAACAAATGTTATACAGGCTGAGAAGAGTCTTAAAAAATATCAAACT	1084
Qy	781	GCAGATGGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAGCAACAAGGA	840
Db	1085	GCAGATGGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAGCAACAAGGA	1144
Qy	841	AGTAGTATTGTTAGTATTTCGTGTTTCCATGCTGTTTAGATGATGCTGAAATCAGACCGCT	900
Db	1145	AGTAGTATTGTTAGTATTTCGTGTTTCCATGCTGTTTAGATGATGCTGAAATCAGACCGCT	1204
Qy	901	TCCATTTTGATGTCTGGGTTTTCGTGCAGATGATTCATGTTCAATACGAAAAATCCTGAT	960
Db	1205	TCCATTTTGATGTCTGGGTTTTCGTGCAGATGATTCATGTTCAATACGAAAAATCCTGAT	1264
Qy	961	TCTCAAGTGCCTCCAAACAGGAGCTCGACACAAAGCTAGAGCAGCGCAAGCCGCTGGAGT	1020
Db	1265	TCTCAAGTGCCTCCAAACAGGAGCTCGACACAAAGCTAGAGCAGCGCAAGCCGCTGGAGT	1324
Qy	1021	GACAGTGTCTGTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTTAGGTAAA	1080
Db	1325	GACAGTGTCTGTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTTAGGTAAA	1384
Qy	1081	GCTGGGCAACAAACGGGCATCTCAATGCTTTAGGACAGATCGCTCTTGCTGCTGTTGG	1140
Db	1385	GCTGGGCAACAAACGGGCATCTCAATGCTTTAGGACAGATCGCTCTTGCTGCTGTTGG	1444
Qy	1141	AGCGCAGGAGT	1151
Db	1445	AGCGCAGGAGT	1455

QY	923	GTCAATGATTTCACATGTTTCAATAAGCGAAATATCTTGATTCTCAAGCTGCCCAACAGGAGC	982
Db	105438	GATCTTAGTCGAACAATTTTAATGTAAACAATCTCTGCACACAGCTAAGAAGCTTACAACTTA	105497
QY	983	TCCGAGCACAAAGCTAGACGACGAAAAGCCGGCTGGAGATGACAGTCTGCTGCAGCGCTGG	1042
Db	105498	TGGAGGCTCAGCTGA--CTGCGATGTCAGATCAACTGGTTGGTGGGATGCGGAGCTCC	105554
QY	1043	CAGATGCTCAGAAAGCTTTAGAAGCGGCTCTAGGTTAAAGCTGGGCAACAAC-----	1093
Db	105555	CAGCGAAATACAAGCAATCAAAGATGCTCTTGCACAAAGCTTTGAAAACAACCATCAGCAG	105614
QY	1094	AGGCACTACTCAATGCTTTTAGACAGATCGCTTCTGCTGCTGTTCTGAGCGCAGGATTC	1153
Db	105615	ATGGTTTAGCTACAGCTATGGACAAGTGGCTTTTGCAGCTCCCAAGTTGGAGGAGGCT	105674
QY	1154	CTCCGGCTGCAACGAAGTTCTTATAGGGTCATCTGTAAAAACAGCTTTTACAAGACTCAAAAT	1213
Db	105675	CCGAGGAAACAGCTGGCACTGTCAGATGAATGTAAAAACAGCTTTTACAAGACAGCGTTTT	105734
QY	1214	CTACAGGTTCT-----GATTATAAAAACAGATATCAGCAGGTTATGATGCTTACAAT	1267
Db	105735	CTTCGACTCTTCCAAGCTCTTATGCAGACAGCATTTCCGATGGATATTCTGCTTACAAA	105794
QY	1268	CCATCAATGATGCCTATGGTAGGGCACCAAAATGATGCGACTCGTGAATGTATAAACAAATG	1327
Db	105795	CAC'TGAATCTTTATATTCOGAAGCAGAGCGGCGTGCAGT'CAGCT'ATTAGTCAAATCG	105854
QY	1328	TAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGCTCGAGGACCG	1387
Db	105855	CAAAATCCCGCCTTTCCAGAAAGCGTTTCTCGTTCTGGCATAGAAAAGTCAAGACGCGAGTG	105914
QY	1388	AAAAACAGATCAAGCCCTCCCTAGGGTGATTTCTGGCAATAGCAGAACTCTTGAGAGTG	1447
Db	105915	CAGATGCTAGCCAAAGACGACGAAACTATTGTCCAGAGATAGCCAAACGTTAGTGATG	105974
QY	1448	TCTATAGTCAAGTTTCGSCACTACAATCTGTAATGCAGATCACTCAGTGGAACTCCTCAAG	1507
Db	105975	TATATAGCCGCTTACAGGTTCTGGATCTTTTGATGCTACGATTCTGAGCAATCCGCAAG	106034
QY	1508	CGAATAATGAGGAGATCAGACAAAGACTTACATCGGAGTGCACAAGCTCCACAGTTTG	1567
Db	106035	CAAAATCAAGAAAGATATATGAGAAAGCTCACGGCACTATTAGCAAAAGCTCCACAATTTG	106094
QY	1568	GCTATCCTTATGTGCAACTTTCTAATGACTCTACACAGAAAGTTTATAGCTAAATTAGAAA	1627
Db	106095	GGTATCCGTGCTGTT'CAGAA'TTCTGTGGATAGCTTGCAGAGTTTCTCTGCAAAATTGGAAA	106154
QY	1628	GTTTGTGTTGCTGAAGGATCTTAGGACAGCAGCTGAAATAAAGACATTTCTCTTTGAAACGA	1687
Db	106155	GAGAGTTTGTGATGGGGAACGTAGTCTCGCAGAATCTCAAGAGAATCGGTTTAGAAAAC	106214
QY	1688	ACTCCTGTTTATTCAGCAGGTGCTGGTCAATATCGGCTCTCTATATTCGTTATCTCC	1747
Db	106215	AGCCGCTTCAITTCACAGGTGTGTAACATGTCTTCTATCTCTCTGGTATCTTT	106274
QY	1748	AATAAC	1753
Db	106275	CTTAAC	106280

RESULT 11

RESUL II
AAZ01425 08

WP Sequence split into 11 fragments LOCUS AAZ01425 Accession Aaz01425
Continuation (9 of 11) of AAZ01425 from base 800001 (Complete genome sequence of Chlamy

	Query Match	10.1%	Score 187.6	DB 2	Length 110000
	Best Local Similarity	53.0%	Pred. No. 1.2e-40		
	Matches 480	Conservative	0	Mismatches 408	Indels 18
	Gaps	3			
Qy	863	TTTCCATGCTGTAGATGCTGAAATGACGCGCTCCATTGTGTCGGTTTC	922		
Db	105378	TTTCTCTTGTCTTCATGTTAGACAATGAAATGGCAGCGATTGAATGCAAGGTTTTC	105437		

Query Match 10.1%; Score 187.6; DB 2; Length 110000;
Best Local Similarity 53.0%; Pred. No. 1.2e-40;
Matches 480; Conservative 0; Mismatches 408; Indels 18;

WP AA201425_07 700001 810000
WP AA201425_08 800001 910000
WP AA201425_09 900001 1010000
WP AA201425_10 1000001 1038602

Query Match 10.1%; Score 187.6; DB 2; Length 110000;
Best Local Similarity 53.0%; Pred. No. 1.2e-40;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY 863 TTTCATGCTGTAGATGCTGAAATAGACCGCTTCATTTTGTGATGCTCTGGGTTTC 922
Db |||||
QY 5378 TTTCCTGTTGCTGTGATGATGAGCAATGAATGCGAGGATGCAATGCAAGGTTTC 5437
Db |||||
QY 923 GTGAGATGATTCATGTTCAATACGGAATCTGATCTCAAGCTGCCCAACAGGAGC 982
Db |||||
QY 5438 GATCTAATGCTGCAACAATTAATGTAAACAATCTCTCAACAGCTAAAGAGCTACAACTA 5497
Db |||||
QY 983 TCGCAGCAAGCTAGACGAGCAAGCGCTGAGATGACATGCTGTGAGCGCTG 1042
Db |||||
QY 5498 TGGAGGCTCAGTGA---CTGCGATGTCAGATCAACTGGTGGTGGGATGGCAGCTCC 5554
Db |||||
QY 1043 CAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGTAAAGCTGGGCAACAC----- 1093
Db |||||
QY 5555 CAGCCGAATATACAAGCAATCAAGATGCTCTTGGCAAGCTTTGAAACAACCATCAGCAG 5614
Db |||||
QY 1094 AGGCGATCTCAATGCTTTAGACAGATGCTCTCTGCTGCTGTGTGAGCGGAGGATTC 1153
Db |||||
QY 5615 ATGGTTAGCTACAGCTATGGGCAAGTGGCTTTTGCAGCTGCCAAGGTTGGAGGAGGCT 5674
Db |||||
QY 1154 CTCCCGCTGCAAGCTCTATAGGCTCATCTGTAACAGCTTTTACAAGACTCAAAAT 1213
Db |||||
QY 5675 CCGCAGGACAGCTGCACTGCTCCAGATGAATGTAACAGCTTTTACAAGACAGCTTTT 5734
Db |||||
QY 1214 CTACAGGTTCT-----GATTATPAAACACAGATATCAGCAGGTTATGATCTTCAAAAT 1267
Db |||||
QY 5735 CTTCGACTCTTCAAGCTCTTATGACAGCAGCACTTTCCGATGATATCTGCTTACAAA 5794
Db |||||
QY 1268 CCATCAATGATGCTATGTTAGGCGACGAATGATGCGACTGCTGATGATGATAACAATG 1327
Db |||||
QY 5795 CACTGAACTCTTTATATCCGAAAGCAGAAAGCGGCTGCGATGAGCTATTAGTCAAACTG 5854
Db |||||
QY 1328 TAAGTACCCCGCTCTCACAGCATCGTTCTTAGACAGCAAGCTCGAGACCGAG 1387
Db |||||
QY 5855 CAAATCCCGGCTTTCCAGAGGGTTCTGCTTCTGGCATAGAAAGTCAAGACGAGTG 5914
Db |||||
QY 1388 AAAAAACAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATAGCAGAACTCTTGGAGATG 1447
Db |||||
QY 5915 CAGATGCTAGCCAAAGACGAGCAGCAAACTATTGTGATGATAGCCAAAGCTTAGGTGATG 5974
Db |||||
QY 1448 TCTATAGTCAAGTTTCGGCACTACATCTGTAATGAGATCACTAGTCGATCTCTCAAG 1507
Db |||||
QY 5975 TATATAGCCGCTTACAGGTTCTGGATCTTTGATGCTACGATGCTGAGCAATCCGCAAG 6034
Db |||||
QY 1508 CGAATANTGAGGAGATCAGACAAAAGCTTACATCGGCACTGACAAAGCTCCACAGTTTG 1567
Db |||||
QY 6035 CAAATCAAGACAGATATTGCAAGAGCTCAGCGCATCTATTAGCAAGCTCCAGATTTG 6094
Db |||||
QY 1568 GCTATCTTATGCGAATTTCTAATGACTTACACAGAGTTTCAATAGTAATTTAGAAA 1627
Db |||||
QY 6095 GGTATCTGCTCTTTCAGAAATCTGTGGATGAGCTTGCAGAGTTTGTGTCAGCAATTTG 6154
Db |||||
QY 1628 GTTGTGTTGCTCAAGATCTAGACAGCAGCTGGAATTAAGCACTTTCTTTTGAACA 1687
Db |||||
QY 6155 GAGAGTTTGTGATGGGAACTGATGCTCGCAGATCTCAAGAGAAATCGTTTGAAGAAC 6214
Db |||||
QY 1688 ACTCCTGTTTATTCAGCAGGTTGGTCAATATCGGCTCTCTATATCTGTTGTTATCTCC 1747
Db |||||
QY 6215 ASCCCGCTTCTTCAACAGGTTGTTGTAACAATTTGCTTCTCTATCTCTGTTTATCTTT 1747
Db |||||
QY 1748 AATAAC 1753
Db |||||
QY 6275 CTTAAC 6280

RESULT 12
AAA63292
XX AAA63292 standard; DNA; 1537 BP.
AC AAA63292;
XX
DT 02-FEB-2001 (first entry)
XX
XX C. trachomatis LGV II clone 12-G3-83 coding sequence.
XX
XX Chlamydial infection; sexually transmitted disease;
XX Chlamydia trachomatis.
XX Chlamydia trachomatis.
XX WO200034483-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 99WO-US029012.
XX
XX 08-DEC-1998; 98US-00208277.
XX 08-APR-1999; 99US-00288594.
XX 01-OCT-1999; 99US-00410568.
XX 22-OCT-1999; 99US-00426571.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX comprises immunogenic portion of Chlamydia antigen, which comprises amino
XX acid sequence encoded by polynucleotide sequence.
XX
XX Claim 1; Page 128-129; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
XX proteins encoded by the nucleic acid sequences. The encoded proteins
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX proteins are useful for the serodiagnosis and treatment of Chlamydia
XX infection. Chlamydiae are intracellular bacterial pathogens that are
XX responsible for a wide variety of human infections. C. trachomatis
XX infection is one of the most common sexually transmitted diseases and can
XX lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX and infertility. Trachoma due to ocular infection with C. trachomatis is
XX the leading cause of preventable blindness worldwide. C. pneumonia is a
XX major cause of acute respiratory tract infections in humans and is also
XX thought to play a role in the pathogenesis of atherosclerosis and
XX coronary heart disease. The present invention is a nucleic acid sequence
XX isolated in the present invention

SQ Sequence 1537 BP; 459 A; 307 C; 357 G; 414 T; 0 U; 0 Other;

Query Match 10.1%; Score 187.2; DB 3; Length 1537;
Best Local Similarity 53.0%; Pred. No. 1.9e-41;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY 863 TTTCATGCTGTAGATGCTGAAATAGACCGCTTCATTTTGTGATGCTCTGGGTTTC 922
Db |||||
QY 186 TTTCCTGTTGCTGTGATGATGAGCAATGAATGCGAGGATGCAATGCAAGGTTTC 245
Db |||||
QY 923 GTGAGATGATTCATGTTCAATACGGAATCTGATTTCTCAAGCTGCCCAACAGGAGC 982
Db |||||
QY 246 GATCTAATGCTGCAACAATTTAATGTAAACAATCTCTCAACAGCTAAAGAGCTACAACTA 305
Db |||||
QY 983 TCGCAGCAAGCTAGACGAGCAAGCGCTGAGATGACAGTGTGCTGACGCTGG 1042
Db |||||
QY 306 TGGAGGCTCAGTGA---CTGCGATGTCAGATCAACTGTTGGTGGGATGGCAGCTCC 362
Db |||||

Db 723 CAGATGCTAGCCAAAGAGCAGCAGAACTATTGTGACAGATAGCCAAAGTTAGGTGATG 782
QY 1448 TCTATAGTCAAGTTTCGGCAGCTACAAATCTGTAAATGAGATCACTCAGTCGAATCTCTCAAG 1507
Db 783 TATATAGCGGCTTACAGGTTCTGGATTCTTTGATGTTCTACGATTGTGAGCAATCCGCAAG 842
QY 1508 CCAATAATGAGGAGATCAGACAAAGCTTACATCGSCAGTGACAAAGCCTCCACAGTTTG 1567
Db 843 CAAATCAAGAAGATTATGCAAGAGCTCAGCGCATCTAATTAGCAAGCTCCACAAATTG 902
QY 1568 GCTATCCTTATGTGCAACTTTCTAATGACTCTACACAGAAAGTTCTATAGTAAATTAGAAA 1627
Db 903 GGTATCCTGCTGTTCAGAACTCTGTGGATAGCTTCGAGAAAGTTTCTGACAAATTGGAA 962
QY 1628 GTTGTGTTGCTGAAGATCTAGGACAGCAGCTGAAATAAAGACACTTTCCTTTGAAACGA 1687
Db 963 GAGAGTTTGTGATGGGGAACGTAGTCTCGCAGAACTCTCAGAGAAATGGTTTGAAGAAC 1022
QY 1688 ACTCCTGTTTATTCAGCAGGTGCTGGTCAATATCGGCTCTATATTTCTGTTATCTCC 1747
Db 1023 AGCCCGCTTCTATTCAACAGGTGTGGTAAACATTGCTTCTATCTCTGTTATCTTT 1082
QY 1748 AATAAC 1753
Db 1083 CTTAAC 1088

RESULT 14

ID ABL92425
XX ABL92425 standard; DNA; 1537 BP.
AC ABL92425;
XX
DT 05-JUN-2002 (first entry)
XX
DE Chlamydia trachomatis DNA sequence SEQ ID NO:57.
XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
KW antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response; gene; ds.
XX

OS Chlamydia trachomatis.
XX

PN WO200208267-A2.
XX

PD 31-JAN-2002.
XX

PF 20-JUL-2001; 2001WO-US023121.
XX

PR 20-JUL-2000; 2000US-00620412.
XX

PR 23-APR-2001; 2001US-00841132.
XX

XX (CORI-) CORIXA CORP.
PA

XX Fling SP, Skeiky YAM, Probst P, Bhatia A;
PI

XX WPI; 2002-179901/23.
DR

XX Novel compositions comprising Chlamydia Cap1 protein and its use in the
PT treatment of Chlamydia infection.
PT

XX Example 1; Page 174; 537pp; English.
PS

CC The present invention describes compositions comprising a Chlamydia Cap1
CC protein and methods for the diagnosis and therapy of Chlamydia infection.
CC Chlamydia DNA and protein sequences from the present invention can have
CC antibacterial and immunostimulant activities, and can be used in
CC vaccines. Compounds from the present invention can be used for eliciting
CC an immune response, specifically stimulating a Chlamydia-specific T-cell
CC response or inhibiting the development of a Chlamydia infection in an
CC animal. Methods from the present invention can be used for detecting the
CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC specific for a Chlamydia protein; and for treatment of a Chlamydia

CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 1537 BP; 459 A; 307 C; 357 G; 414 T; 0 U; 0 Other;

Query Match 10.1%; Score 187.2; DB 6; Length 1537;
Best Local Similarity 53.0%; Pred. No. 1.9e-41;

Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY 863 TTTCATCTCTGTAGATGATGCTGAAAATGAGACGCGTTCCATTTTGTGTCGGTTTC 922
Db 186 TTTCCTCTGTTGATGATGTAGACATGAATGGCAGCGATTGCAATGCAAGTTTC 245
QY 923 GTCAGATGATTCACATGTTCAATAGCGAAATTCCTGATTCCTCAAGCTGCCCAACAGAGC 982
Db 246 GATCTATGATCGAAATTTAATGTAAATCTCTGCAACAGCTAAAGAGGTACAAAGCTA 305
QY 983 TCGCAGACAAAGCTAGACGACGAAAGCGCTGGAGATGACAGTCTGCTCAGCGCTGG 1042
Db 306 TGGAGGCTCAGTGA---CTGCGATGTCAATCACTGGTTGGTGGGATGGCAGCTCC 362
QY 1043 CAGATGCTCAGAAAAGCTTTAGAAGCGCTCTAGGTAAAGCTGGGCAACAAC----- 1093
Db 363 CAGCGAAATACAAAGCAATCAAGATGCTCTTGGCAAGCTTTGAAACAACCATCAGCAG 422
QY 1094 AGGCAATCACTCAATGCTTTAGGACAGATGCTTCTGCTGCTGTTGTGAGCGCAGAGTTC 1153
Db 423 ATGTTTAGCTACAGCTATGGCAAGTGGCTTTTTCAGCTGCCAAGTTGGAGAGGCT 482
QY 1154 CTCGCGCTGCAGCAAGTTCTATAGGTCTATGTAAGACAGCTTTTACAAAGACCTCAAAAT 1213
Db 483 CCGCAGNACAGCTGGGCACTGTCAGATGAATGTAAACACGCTTTAAGACAGGTTTT 542
QY 1214 CTACAGGTTCT-----GATTATAAAACACAGATATACAGAGTTATGATGCTTACAAAT 1267
Db 543 CTTGACTTCTTCCAGCTCTTATGCAGCAGCACCTTCGATGGATATTTCTGTTACAAA 602
QY 1268 CCATCAATGATCGCTATGTTAGGCGCAGAAATGATGCGACTCGTGATGTATAACATG 1327
Db 603 CACTGAACCTTTTATATTCGAAAGCAGAGCGGCTGCACTGATTTAGTCAAACTG 662
QY 1328 TAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGCTCGAGGACCAG 1387
Db 663 CAATCCGCGCTTCCAGAGCGTTTCTCGTTCTGGCATAGAAGTCAAGAGCGCAGTG 722
QY 1388 AAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATPAGCAGAACTCTTGGAGATG 1447
Db 723 CAGATGCTAGCCAAAGAGCAGCAGAAACTATTGTACAGAGATAGCCAAACGTTAGTGATG 782
QY 1448 TCTATAGTCAAGTTTCGGCAGCTACAAATCTGTAAATGACAGATCACTCAGTCGATCTCAAG 1507
Db 783 TATATAGCGGCTTACAGGTTCTGGAATCTTTGATGTTCTAGCAATTTGTGAGCAATCCGCAAG 842
QY 1508 CGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCTCCACAGTTTG 1567
Db 843 CAATCAAGAAGAGATTATGCAAGCTCAGCGCATCTATTAGCAAGCTCCCAAAATTG 902
QY 1568 GCTATCCTTATGTGCAACTTTCTAATGACTCTACACAGAGTTTCAATAGTAAATTAGAAA 1627
Db 903 GGTATCCTGCTGTTTCAGAAATTTCTGTGGATAGCTTTGCAAGAGTTTGTCTGCAAAATTGAAA 962
QY 1628 GTTTGTTTGTGTAAGGATCTAGGACAGCAGCTGAAATAAAGACACTTTCCTTTGAAACGA 1687
Db 963 GAGAGTTTGTGATGGGGAACGTAGTCTCGCAGAACTCTCAGAGAAATGCGTTTAGAAAAA 1022
QY 1688 ACTCCTTGTATTATTCAGCAGGTGCTGGTCAATATCGGCTCTCTATATTTCTGTTATCTCC 1747
Db 1023 AGCCCGCTTCTATTCAACAGGTGTGGTAAACATTGCTTCTATCTCTGTTATCTTT 1082
QY 1748 AATAAC 1753
Db 1083 CTTAAC 1088

RESULT 15

ADD43867
ID ADD43867 standard; DNA; 1944 BP.

XX AC ADD43867;

XX DT 15-JAN-2004 (first entry)

XX DE Chlamydia trachomatis immunogenic DNA sequence, SEQ ID No 162.

XX DX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;

XX KW gene therapy; antibacterial; gene; ds.

XX OS Chlamydia trachomatis.

XX FN WO2003049762-A2.

XX PD 19-JUN-2003.

XX PF 12-DEC-2002; 2002WO-IB005761.

XX PR 12-DEC-2001; 2001GB-00029732.

XX PR 06-AUG-2002; 2002GB-00018233.

XX PR 14-AUG-2002; 2002GB-00018924.

XX PA (CHIR-) CHIRON SPA.

XX PI Grandi G, Ratti G;

XX DR WPI; 2003-532882/50.

XX DR P-PSDB; ADD43866.

XX PT New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.

XX PS Claim 7; SEQ ID NO 162; 164pp; English.

XX CC The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This polynucleotide sequence
CC represents one of the 131 DNA sequences encoding the C. trachomatis
CC proteins with immunogenic properties of the invention.

XX SQ Sequence 1944 BP; 612 A; 391 C; 445 G; 496 T; 0 U; 0 Other;

Query Match 10.0%; Score 186; DB 9; Length 1944;

Best Local Similarity 48.5%; Pred. No. 4.6e-41;

Matches 720; Conservative 0; Mismatches 725; Indels 39; Gaps 6;

QY 302 AAGTTGGCGCGCAATTACAGAAATAGCGAAATATGCTTCGGATAACCAAGGATCTTGTG 361

DB 467 AGGTTATTGATGAAGCTCAATCAGTTAGTTAACTTGAATAACAGAAATCAGACTTTAAAGG 526

QY 362 ACTCTTTAGGTTAACTGACTTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTAG 421

DB 527 AAACCTTTAAACCAACACAGACTCTGCAGATCAGATTCCAGCGATTAAATAGTCAGTTAGAGA 586

QY 422 CAAACAATTAACAACGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCGAG 481

DB 587 TCAACAAAAATTCTGCAGATCAAAATTATCAAGATCTGGAAGGCAAAAACATAAGTTATG 646

QY 482 GGAAACGCCCTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAGA 541

DB 647 AAGCTGTTCTCACTAACCGCAGGAGAGGTTATCAAAAGCTTCTTCTGAAAGCGGAATTAAGT 706

QY 542 TAGAGAAAGATGGAAATGCGATTAGGGATGCATATTTTTCAGGACAGAAACGCTAGTGGAG 601

DB 707 TAGACAAAGCTTTGCACTCTATTGTGATGCTGGGATCAAAAGCGGCTGCACTTTCTTC 766

QY 602 CTGTAGAAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTAAACAGCAA 661

DB 767 AAGCAGCAGCAAAATAATAGCCAGATATATCGCAGCCACGAGAAATTAATTGATGCTG 826

QY 662 TCGCTACTGTCTAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGCAGCTCTTCCAATTC 721

DB 827 CTGAACGCAAGGTAAACGAGTTAAAAACAAGACATACAGGGCTAACCGACTTCGCTTTAG 886

QY 722 TTCAGAAGCGGACAAATGGTATATACAGGCTGAGGAAGATCTTAAATAATCAAACTG 781

DB 887 TGAATAAAGCTGAGGAGCAGATTAGTCAAGCAAAAAGATATTCAAGAGATCAAACTC 945

QY 782 CAGATGTTCTGTATGTTTCCAAATCCAGGAATCAGATTGGAG-----GCTCCAAGCAACA 836

DB 946 --AGTGGTTCGGATATTCTTATCGTTGGTCCGAGTGGGTGAGTCTCCGCAAGGAAGTG 1003

QY 837 AGGAAGTAGTATTGGTAGTAT-----TCGTGTTTCCATGCTGTTAGATGATG 893

DB 1004 CGGTAGGAGCGTTGAAATCTCTTAAACAATTCAAGGAAGAAATTTCTCTTCTGTTGATG 1063

QY 884 CTGAAAAATGAGACCGCTTCCATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 943

DB 1064 TAGCAATGAATGGCAGCGATTCGATGCAATGCAAGTTTTCGATCTATGATGCAAAATTA 1123

QY 944 ATACGGAAAAATCCTGTGATTTCTCAAGCTCCCAACAGGAGCTCGCAGCAACAAGCTAGACAG 1003

DB 1124 ATGTAAACAATCTCTGCAACAGCTAAAGAGCTACAAGCTATGGAGGCTCAGCTGA---CTG 1180

QY 1004 CGAAGCGCTGGAGATGACAGTCTGCTGAGCGCTGGGAGATGCTCAGAAAGCTTTAG 1063

DB 1181 CGATGTAGATCAACTGTTGGTGGGATGGGAGCTCCAGCGGCAAAATACAAAGCAATCA 1240

QY 1064 AAGCGGCTCTAGTAAAGCTGGGCAACAAC-----AGGGCATACTCAATGCTTTAG 1114

DB 1241 AAGATGCTCTTGGCAGACTTTGAAACAACCAATCAACAGATGTTTATAGCTACAGCTATGG 1300

QY 1115 GACGATGCTCTTCTGCTGCTGTTGTGAGCGCAGAGGTTCTCCCGCTGCGAGCAAGTTCTTA 1174

DB 1301 GACAAGTGGCTTTTTCGAGCTGCCAAGGTTGGAGGAGGCTCCGCAAGGAACAGCTGGCACTG 1360

QY 1175 TAGGTCATCTGTAAAAACAGCTTTACAGACCTCAAAATCTACAGGTTCTGA-----TT 1228

DB 1361 TCCAGATGAATGTAACAACAGCTTTACAGCAGGTTTTTCTTCGACTTCTTCCAGCTCTT 1420

QY 1229 ATAAAAACAGATATACGAGGTTTATGATCTTACAAATCCATCAATGATGCTCCATGGTA 1288

DB 1421 ATGCAGCAGCACTTTCCGATGGATATTCTGCTTACAAAAACACTGAACCTCTTTATATCCG 1480

QY 1289 GGGCAGCAAAATGATGCGACTCGTGATGTGATAACAATTAAGTAAGTACCCCGCTCTACAC 1348

DB 1481 AAAGCAGAAAGCGGCTGCGAGCTAGCTATTAGTCAAACTGCAAAATCCCGGCTTTTCCAGAA 1540

QY 1349 GATCCGTTCTTAGAGCAGCAACAGAGCTCGAGACCCAGAAAAACAGATCAAGACCCCTCG 1408

DB 1541 GCGTTTCTGTTCTGGCATAGAAAGTCAAGAGCGAGTGCAGATGCTAGCCCAAGAGCAG 1600

QY 1409 CTAGGGTGATTTCTGGCAATAGCAGAACTCTTTGAGATGCTTATAGTCAAGTTTTCGSCAC 1468

DB 1601 CAGAAACTATTGTTCAGAGATAGCCAAACGTTAGGTGATGATATATAGCCGCTTACAGGTTT 1660

QY 1469 TACAATCTGTANTGCAGATCACTCAGTCTGAATCTCTAGCGCAATTAATGAGAGATCAGAC 1528

DB 1661 TGGATTCTTTGATGTTACGATTGTGAGCAATCCGCAAGTAAATCAAGAGAGATATTATGC 1720

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd..

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:23:54 ; Search time 639.252 Seconds
(without alignments)
10716.954 Million cell updates/sec

Title: US-10-608-559-3
Perfect score: 1852
Sequence: 1 atgagctgcgcagataagct.....gatcttcataataaaga 1852

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1852	100.0	1852	US-09-564-479-3	Sequence 3, Appli
2	1848.8	99.8	2156	US-09-564-479-1	Sequence 1, Appli
3	1836.2	99.1	1230025	US-10-289-762-1	Sequence 1, Appli
4	1748.8	94.4	1956	US-09-841-132-385	Sequence 385, Appl
5	1748.8	94.4	1956	US-09-841-260-63	Sequence 63, Appl
6	1748.8	94.4	1956	US-10-007-693-63	Sequence 63, Appl
7	1748.8	94.4	1956	US-10-312-273-14	Sequence 14, Appl
8	1161.8	62.7	2238	US-09-564-479-7	Sequence 7, Appli
9	1151	62.1	1456	US-09-564-479-5	Sequence 5, Appli
10	187.2	10.1	1637	US-09-841-132-57	Sequence 57, Appl
11	185.4	10.0	1941	US-09-841-132-412	Sequence 412, Appl
12	184	9.9	1171	US-09-841-260-13	Sequence 13, Appl
13	184	9.9	1171	US-10-007-693-13	Sequence 13, Appl
14	184	9.9	1834	US-09-841-260-38	Sequence 38, Appl
15	184	9.9	1834	US-10-007-693-38	Sequence 38, Appl

16	183	9.9	1983	10	US-09-841-260-28	Sequence 28, Appl
17	183	9.9	1983	13	US-10-007-693-28	Sequence 28, Appl
18	49	2.6	1860	12	US-10-282-122A-37518	Sequence 37518, A
19	47	2.5	9373	12	US-10-221-613-178	Sequence 178, App
20	47	2.5	9373	14	US-10-311-455-1264	Sequence 1264, Ap
21	45.2	2.4	2404	15	US-10-104-047-1740	Sequence 1740, Ap
22	45.2	2.4	3439	15	US-10-161-827-53	Sequence 53, Appl
23	44.4	2.4	2499	9	US-09-842-552-92	Sequence 92, Appl
24	44.4	2.4	4997	12	US-10-282-122A-35506	Sequence 35506, A
25	43.8	2.4	702	15	US-10-027-632-45880	Sequence 45880, A
26	43.8	2.4	702	15	US-10-027-632-45881	Sequence 45881, A
27	42.6	2.3	1431	12	US-10-424-599-127608	Sequence 127608,
28	42.2	2.3	897	12	US-10-424-599-127609	Sequence 127609,
29	41.4	2.2	1398	12	US-10-282-122A-15784	Sequence 15784, A
30	41.2	2.2	1237	12	US-10-412-699B-453	Sequence 453, App
31	41.2	2.2	1237	14	US-10-295-403-49	Sequence 49, Appl
32	41.2	2.2	1368	12	US-10-425-114-14565	Sequence 14565, A
33	41.2	2.2	1737	14	US-10-032-585-6793	Sequence 6793, Ap
34	40.6	2.2	618	15	US-10-027-632-269245	Sequence 269245,
35	40.6	2.2	3249	12	US-10-424-599-139567	Sequence 139567,
36	40.4	2.2	512	9	US-09-864-761-25347	Sequence 25347, A
37	40.4	2.2	575	9	US-09-864-761-8628	Sequence 8628, Ap
38	40.4	2.2	3137	14	US-10-037-270-714	Sequence 714, App
39	40.4	2.2	3137	15	US-10-117-722-714	Sequence 714, App
40	40.4	2.2	12541	15	US-10-074-024-524	Sequence 524, App
41	40.4	2.2	12541	15	US-10-074-024-525	Sequence 525, App
42	40	2.2	7311	9	US-09-815-242-4779	Sequence 4779, Ap
43	40	2.2	18846	9	US-09-815-242-8898	Sequence 8898, Ap
44	39.8	2.1	401	9	US-09-815-242-2951	Sequence 2951, Ap
45	39.8	2.1	401	12	US-10-282-122A-5515	Sequence 5515, Ap

ALIGNMENTS

RESULT 1

US-09-564-479-3
; Sequence 3, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: 032931/0230
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1749)
US-09-564-479-3

Query Match 100.0%; Score 1852; DB 10; Length 1852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAGCTCGGAGATAAGCTGGGTATTCCTTCTAGTAACAGCTCGTCTTCTACTACGAGA	60
Db	1	ATGAGCTCGGAGATAAGCTGGGTATTCCTTCTAGTAACAGCTCGTCTTCTACTACGAGA	60
Qy	61	TCGACAGCTGGACTCAACGACGACCGACCTACGCTCTCTCCACCCAGTTTGTAT	120

Db 1141 AGCGCAGAGTTCTCCCGCTGCGCAAGTTCTATAGGTCATCTGTAAACACGCTTTAC 1200
Qy 1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGCT 1260
Db 1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGCT 1260
Qy 1261 TACAAATCCATCAATGATGCTATGTTAGGCGCAGAAATGATGCGATCTCGTATGATGATA 1320
Db 1261 TACAAATCCATCAATGATGCTATGTTAGGCGCAGAAATGATGCGATCTCGTATGATGATA 1320
Qy 1321 AACAAATGATGATGCTATGTTAGGCGCAGAAATGATGCGATCTCGTATGATGATA 1380
Db 1321 AACAAATGATGATGCTATGTTAGGCGCAGAAATGATGCGATCTCGTATGATGATA 1380
Qy 1381 GGACCAAGAAACACAGATCAAGCCCTGCTAGGTCGATTTCTGCAATACAGAACTCTT 1440
Db 1381 GGACCAAGAAACACAGATCAAGCCCTGCTAGGTCGATTTCTGCAATACAGAACTCTT 1440
Qy 1441 GGAGATCTTATAGTCAAGTTTCGGCACTACAAATCTGTAATGAGATCACTCAGTCGAAT 1500
Db 1441 GGAGATCTTATAGTCAAGTTTCGGCACTACAAATCTGTAATGAGATCACTCAGTCGAAT 1500
Qy 1501 CCTCAAGCAATATGAGGAGATCAGACAAAGCTTACATCGGCAAGTACAAAGCTTCCA 1560
Db 1501 CCTCAAGCAATATGAGGAGATCAGACAAAGCTTACATCGGCAAGTACAAAGCTTCCA 1560
Qy 1561 CAGTTTGGCTATCTTATGTCACCTTTCTTAATGACCTCTACACAGAACTTCTATGCTTAA 1620
Db 1561 CAGTTTGGCTATCTTATGTCACCTTTCTTAATGACCTCTACACAGAACTTCTATGCTTAA 1620
Qy 1621 TTAGAAAGTTTGTGCTGAGGATCTAGACACAGCTGAAATATAAGCACTTTCCTTT 1680
Db 1621 TTAGAAAGTTTGTGCTGAGGATCTAGACACAGCTGAAATATAAGCACTTTCCTTT 1680
Qy 1681 GAAACGAACTCTTGTGTTTATGTCAGCAGTGTGCTGCTCAATATCGGCTCTTATATCTGCT 1740
Db 1681 GAAACGAACTCTTGTGTTTATGTCAGCAGTGTGCTGCTCAATATCGGCTCTTATATCTGCT 1740
Qy 1741 TATCTCAATATAACACACCTAAAGTTTCTGTTTGGAGAGATTTATGCTGCTTGGTAAGG 1800
Db 1741 TATCTCAATATAACACACCTAAAGTTTCTGTTTGGAGAGATTTATGCTGCTTGGTAAGG 1800
Qy 1801 CCTTTGTGAGGCTTACCAACACACTAGAACGATCTTCAATATAAAGAA 1852
Db 1801 CCTTTGTGAGGCTTACCAACACACTAGAACGATCTTCAATATAAAGAA 1852

RESULT 2

US-09-564-479-1

; Sequence 1, Application US/09564479

; Publication No. US2003009573A1

; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: OMEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND

; FILE OF INVENTION: US/09564, 479

; FILE REFERENCE: 032931/0230

; CURRENT APPLICATION NUMBER: US/09564, 479

; PRIOR FILING DATE: 2000-05-03

; PRIOR APPLICATION NUMBER: 60/132,270

; PRIOR FILING DATE: 1999-05-03

; PRIOR APPLICATION NUMBER: 60/141,276

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2156

; TYPE: DNA

; ORGANISM: Chlamydia pneumoniae

; FEATURE:

; NAME/KEY: CDS

LOCATION: (101)..(2053)
US-09-564-479-1

Query Match 99.8%; Score 1848.8; DB 10; Length 2156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGAGTCTGGCAGATAAGCTGGGTATTTGCTTTAGTAAACAGCTCGTCTTACTAGCAGA	60
DB	305	ATGAGTCTGGCAGATAAGCTGGGTATTTGCTTTAGTAAACAGCTCGTCTTACTAGCAGA	364
QY	61	TTCTGACAGCTGGACTCAACGACAGCAGCGACCTACGCTCCCTCCACCCAGCTTTGAT	120
DB	365	TTCTGACAGCTGGACTCAACGACAGCAGCGACCTACGCTCCCTCCACCCAGCTTTGAT	424
QY	121	GATTATAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT	180
DB	425	GATTATAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT	484
QY	181	GACATACAGCTGCTTTGGTGGAGCTCCAGGATGCTGTCACTAAATAAAGGATACAGCG	240
DB	485	GACATACAGCTGCTTTGGTGGAGCTCCAGGATGCTGTCACTAAATAAAGGATACAGCG	544
QY	241	GCTACTGATGAGGAAACCCGCAATCGCTGCGAGTGGGAACTAAGAAATGCCGATGCAGTT	300
DB	545	GCTACTGATGAGGAAACCCGCAATCGCTGCGAGTGGGAACTAAGAAATGCCGATGCAGTT	604
QY	301	AAAGTTGGCGCGCAAAATTAAGAAATGCGGATTAACCAAGCGATTTCT	360
DB	605	AAAGTTGGCGCGCAAAATTAAGAAATGCGGATTAACCAAGCGATTTCT	664
QY	361	GACTCTTTAGTAACTGCTTCTTCCGACCTTTACAGGCTGCTTTTCCAAATCTGTA	420
DB	665	GACTCTTTAGTAACTGCTTCTTCCGACCTTTACAGGCTGCTTTTCCAAATCTGTA	724
QY	421	GCAACAATTAACAAAGCAGCTGAGCTTTCTTAAGAGATGCAAGATAACCCAGTAGTCCA	480
DB	725	GCAACAATTAACAAAGCAGCTGAGCTTTCTTAAGAGATGCAAGATAACCCAGTAGTCCA	784
QY	481	GGGAAACCCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACACGACACAG	540
DB	785	GGGAAACCCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACACGACACAG	844
QY	541	ATAGAGAAAGATGGAATTCGGATTTAGGATGATGATTTTTCAGGACAGAACCTAGTGA	600
DB	845	ATAGAGAAAGATGGAATTCGGATTTAGGATGATGATTTTTCAGGACAGAACCTAGTGA	904
QY	601	GCTGTAGAAATGCTAAATCTAATAACAGTAAAGCAATAGATTCAGTAAAGCAGCA	660
DB	905	GCTGTAGAAATGCTAAATCTAATAACAGTAAAGCAATAGATTCAGTAAAGCAGCA	964
QY	661	ATCGTACTGCTTAAGACACAAATAGCTGAGCTCAGAAAAAGTTCCCGACTCTCCAAT	720
DB	965	ATCGTACTGCTTAAGACACAAATAGCTGAGCTCAGAAAAAGTTCCCGACTCTCCAAT	1024
QY	721	CTTCAAGAAAGCGAAACAAATGTAATACAGCTGAGAAAGATCTTAAATAATCAAACT	780
DB	1025	CTTCAAGAAAGCGAAACAAATGTAATACAGCTGAGAAAGATCTTAAATAATCAAACT	1084
QY	781	GCAGATGCTCTGATGTTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACAGGA	840
DB	1085	GCAGATGCTCTGATGTTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACAGGA	1144
QY	841	AGTAGTATTTGGTAGTATTTGGTGGTTCGATGCTTTAGATGATGCTGAAAAATGAGACGCT	900
DB	1145	AGTAGTATTTGGTAGTATTTGGTGGTTCGATGCTTTAGATGATGCTGAAAAATGAGACGCT	1204
QY	901	TCCATTTTGTCTCGGTTTGGTTCGATGATGATTCATGTTCAATACGAAAAATCTCTGAT	960
DB	1205	TCCATTTTGTCTCGGTTTGGTTCGATGATGATTCATGTTCAATACGAAAAATCTCTGAT	1264
QY	961	TCTCAAGCTGCCAACAGGAGCTCCGACGACAAGCTAGACGAGCAAGCCGCTGGAGAT	1020

DB	1265	TCTCAAGCTGCCAACACAGGAGCTCGACGACCAAGCTAGACGACGAAAGCCGCTGGAGAT	1324
QY	1021	GACAGTGTCTGTGACAGCGCTGSCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA	1080
DB	1325	GACAGTGTCTGTGACAGCGCTGSCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA	1384
QY	1081	GCTGGGCAACAAACAGGGCATACTCAATGCTTTAGACAGATCGCTTCTGCTCTGTGTG	1140
DB	1385	GCTGGGCAACAAACAGGGCATACTCAATGCTTTAGACAGATCGCTTCTGCTCTGTGTG	1444
QY	1141	AGCGAGGAGTTCTTCCGCTGACAGAAAGTTCTATAGGTTATCTGTGTAAGACAGCTTTAC	1200
DB	1445	AGCGAGGAGTTCTTCCGCTGACAGAAAGTTCTATAGGTTATCTGTGTAAGACAGCTTTAC	1504
QY	1201	AAGACCTCAAAATCTACAGGTTCTGATTATATAAACAACAGATATCAGAGGTTATGATCT	1260
DB	1505	AAGACCTCAAAATCTACAGGTTCTGATTATATAAACAACAGATATCAGAGGTTATGATCT	1564
QY	1261	TACAAATCAATCAATGATGCTTATGTTAGGGCAAGAAATGATGCGACTCGTATGATGATA	1320
DB	1565	TACAAATCAATCAATGATGCTTATGTTAGGGCAAGAAATGATGCGACTCGTATGATGATA	1624
QY	1321	AACATGTAAGTACCCCGCTCTCAGAGTCCGTTCTAGAGCAGCAACAGAACTCGA	1380
DB	1625	AACATGTAAGTACCCCGCTCTCAGAGTCCGTTCTAGAGCAGCAACAGAACTCGA	1684
QY	1381	GGACAGAAAAAACAAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATAGCAGAACTCTT	1440
DB	1685	GGACAGAAAAAACAAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATAGCAGAACTCTT	1744
QY	1441	GGAGATGCTATGATCAAGTTTCCGCACTACAATCTGTAATGCGAGATCATCCAGTCAAT	1500
DB	1745	GGAGATGCTATGATCAAGTTTCCGCACTACAATCTGTAATGCGAGATCATCCAGTCAAT	1804
QY	1501	CCTCAAGCGAATAATAGAGAGATCAGACAAAGCTTACATCGGAGTCAACAAAGCTCCA	1560
DB	1805	CCTCAAGCGAATAATAGAGAGATCAGACAAAGCTTACATCGGAGTCAACAAAGCTCCA	1864
QY	1561	CAGTTTGGCTATCTCTTATGTCGCACTTTCTAATGACTCAACAGAACTTCATAGCTAAA	1620
DB	1865	CAGTTTGGCTATCTCTTATGTCGCACTTTCTAATGACTCAACAGAACTTCATAGCTAAA	1924
QY	1621	TTAGAAAGTTTGTGCTGAAAGGATCTAGGACAGAGCTGAAATAAAGCACTTTCCCTTT	1680
DB	1925	TTAGAAAGTTTGTGCTGAAAGGATCTAGGACAGAGCTGAAATAAAGCACTTTCCCTTT	1984
QY	1681	GAAACGAACTCTCTGTTTATTTTACAGAGTGTGTTTGGAGAGATTTATGTTGTTGTAAGG	1740
DB	1985	GAAACGAACTCTCTGTTTATTTTACAGAGTGTGTTTGGAGAGATTTATGTTGTTGTAAGG	2044
QY	1741	TATCTTCAATAAACACACCTTAAGTGTGTTTGGAGAGATTTATGTTGTTGTAAGG	1800
DB	2045	TATCTTCAATAAACACACCTTAAGTGTGTTTGGAGAGATTTATGTTGTTGTAAGG	2104
QY	1801	CCTTTGTTGAGGCTTACCAACACACACTAGAACGATCTTCAATAAATAAAGA	1852
DB	2105	CCTTTGTTGAGGCTTACCAACACACACTAGAACGATCTTCAATAAATAAAGA	2156

RESULT 3

US-10-289-762-1/c
; Sequence 1, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289/762
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1

Handwritten signature

```
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (450001)..(465000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (510001)..(525000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (525001)..(540000)
```

/	/	OTHER INFORMATION: n=a or c or g or t	Db	828034	GACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA	827975
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	421	GCAAA CAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTAGTCCCA	480
/	/	LOCATION: (540001)..(555000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827974	GCAAA CAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTAGTCCCA	827915
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	481	GGGAAAA CGCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGACAG	540
/	/	LOCATION: (555001)..(570000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827914	GGGAAAA CGCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGACAG	827855
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	541	ATAGAGAAAGATGGAAATGCGAATTAGGGATGATATTTTGACGACAGAACTAGTGA	600
/	/	LOCATION: (570001)..(585000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827854	ATAGAGAAAGATGGAAATGCGAATTAGGGATGATATTTTGACGACAGAACTAGTGA	827795
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	601	GCTGTAGAAATGCTAAATCTAATAACAGATTAAGCAACATAGATTCAGCTAAAGCAGCA	660
/	/	LOCATION: (585001)..(600000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827794	GCTGTAGAAATGCTAAATCTAATAACAGATTAAGCAACATAGATTCAGCTAAAGCAGCA	827735
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	661	ATCGCTACTGCTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGACTCTCCAAAT	720
/	/	LOCATION: (600001)..(615000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827734	ATCGCTACTGCTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGACTCTCCAAAT	827675
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	721	CTTCAAGAAAGCGGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT	780
/	/	LOCATION: (615001)..(630000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827674	CTTCAAGAAAGCGGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT	827615
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	781	GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACAGGA	840
/	/	LOCATION: (630001)..(645000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827614	GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACAGGA	827555
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	841	AGTAGTATTGGTAGTATTTCGTTGTTTCCATGCTGTAGATGCTGAAATAGAGACCGCT	900
/	/	LOCATION: (645001)..(660000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827554	AGTAGTATTGGTAGTATTTCGTTGTTTCCATGCTGTAGATGCTGAAATAGAGACCGCT	827495
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	901	TCCATTTTGATGCTCTGGGTTTCGTCAGATGATTCACATGTTCAATACGAAATCTCGAT	960
/	/	LOCATION: (660001)..(675000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827494	TCCATTTTGATGCTCTGGGTTTCGTCAGATGATTCACATGTTCAATACGAAATCTCGAT	827435
/	/	FEATURE:				
/	/	NAME/KEY: misc_feature	QY	961	TCTCAAGCTGCCCAACAGGAGCTCCAGCAACAGCTAGCAGCGCAAGACCGCTCGAGAT	1020
/	/	LOCATION: (675001)..(690000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827434	TCTCAAGCTGCCCAACAGGAGCTCCAGCAACAGCTAGCAGCGCAAGACCGCTCGAGAT	827375
/	/	FEATURE:				
/	/	Query Match	QY	1021	GACAGTGTCTGTCAGCGCTGSCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGTAAA	1080
/	/	Best Local Similarity	Db	827374	GACAGTGTCTGTCAGCGCTGSCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGTAAA	827315
/	/	Matches 1849; Conservative 0; Mismatches 3; Indels 1; Gaps 1;				
/	/	1	QY	1081	GCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATGCTTCTGCTGCTG-TTGT	1139
/	/	ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	Db	827314	GCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATGCTTCTGCTGCTGTTGT	827255
/	/	828394				
/	/	61	QY	1140	GAGCGCAGAGTTCTCTCCGCTGAGCAAGTCTATAGGGTCACTCTGTAAACACAGTTTA	1199
/	/	TCTGAGAGCTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCAGCTTTGAT	Db	827254	GAGCGCAGAGTTCTCTCCGCTGAGCAAGTCTATAGGGTCACTCTGTAAACACAGTTTA	827195
/	/	828334				
/	/	121	QY	1203	CAAGACCTCAAAATCTCAGAGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGC	1259
/	/	GATTATAAGACTCAACGCGAAACAGCTTACGATCTATCTTTAGCTCAACATCACTAGCT	Db	827194	CAAGACCTCAAAATCTCAGAGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGC	827135
/	/	828274				
/	/	181	QY	1260	TTACAAATCCATCAATGATGCTATGGTAGGCGCAGAAATGATGCGATCGTGATGTGAT	1319
/	/	GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGCTCACTAATATAAGGATACAGCG	Db	827134	TTACAAATCCATCAATGATGCTATGGTAGGCGCAGAAATGATGCGATCGTGATGTGAT	827075
/	/	828214				
/	/	241	QY	1320	AAACAAATGATGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCAGTTCG	1379
/	/	GCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCAGTTCG	Db	827074	AAACAAATGATGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCAGTTCG	827015
/	/	828154				
/	/	301	QY	1380	AGGACCCAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCT	1439
/	/	AAAGTTGGCGCGCAAAATTACAGAAATAGCTTCGATTAACCAACGATTTCT	Db	827014	AGGACCCAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCT	826955
/	/	828094				
/	/	361	QY	1440	TGGAGATCTCTATAGTCAAGTTTCGGCACTACAAATCTGTAATGCAGATCACTCACTGCGAA	1499
/	/	GACTCTTTAGGTAAACTGACTTCTCTCGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA	Db	826954	TGGAGATCTCTATAGTCAAGTTTCGGCACTACAAATCTGTAATGCAGATCACTCACTGCGAA	826895
/	/					

QY 1500 TCCTCAAGCGGAATATAGGAGATCAGACAAAAGCTTACATCGGAGTGACAAAGCCTCC 1559
Db 826894 TCCTCAAGCGGAATATAGGAGATCAGACAAAAGCTTACATCGGAGTGACAAAGCCTCC 826895
QY 1560 ACAGTTTGGCTATCTTATGTGCAACTTTCTAATGACTCTACACAGAAGTTTCATAGCTAA 1619
Db 826894 ACAGTTTGGCTATCTTATGTGCAACTTTCTAATGACTCTACACAGAAGTTTCATAGCTAA 826775
QY 1620 ATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGCTGAATATAAAGCACTTTCCTT 1679
Db 826774 ATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGCTGAATATAAAGCACTTTCCTT 826715
QY 1680 TGAACGAACTCTCTTGTATTTACGAGGCTGCTGCTAATATGCGCTCTCTATATCTGG 1739
Db 826714 TGAACGAACTCTCTTGTATTTACGAGGCTGCTGCTAATATGCGCTCTCTATATCTGG 826655
QY 1740 TTATCTCAATAACAAACCTTAAGTGTGCTTTGGAGAGATTTATATGCTTTGGTAAG 1799
Db 826654 TTATCTCAATAACAAACCTTAAGTGTGCTTTGGAGAGATTTATATGCTTTGGTAAG 826595
QY 1800 GCCTTTGTTAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 1852
Db 826594 GCCTTTGTTAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 826542

RESULT 4
US-09-841-132-385
; Sequence 385, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yashir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 385
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-385

Query Match 94.4%; Score 1748.8; DB 9; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAGAGTGGTATGCTTCTAGTAAACAGTCTGCTTCTACTAGCAGA 60
Db 205 ATGAGTCTGGCAGATAGAGTGGTATGCTTCTAGTAAACAGTCTGCTTCTACTAGCAGA 264
QY 61 TCTGACAGCTGGACTCAACGACGACGACGACCTAGCTCTCTCCACCCAGCTTTGAT 120
Db 265 TCTGACAGCTGGACTCAACGACGACGACGACCTAGCTCTCTCCACCCAGCTTTGAT 324
QY 121 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCAGTACT 180
Db 325 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCAGTACT 384
QY 181 GACATACAGGCTCTTTGGTAGCTCCAGAGTCTGTCATATATATAAGGATACAGCG 240
Db 385 GACATACAGGCTCTTTGGTAGCTCCAGAGTCTGTCATATATATAAGGATACAGCG 444
QY 241 GCTACTGATGAGGAACCGCAATCTGCGGAGTGGGAACTAGAAATGCGGATGCACTT 300
Db 445 GCTACTGATGAGGAACCGCAATCTGCGGAGTGGGAACTAGAAATGCGGATGCACTT 504
QY 301 AAAGTTGGCGCGCAAAATACAGAATTAGCGAAATAGCTTCGGATAACCAAGCGGATCTT 360

Db 505 AAAGTTGGCGCGCAAAATACAGAATTAGCGAAATATGCTTCGGATATACCAAGCGAATCTTT 564
QY 361 GACTCTTTAGGTAAATGACTGCTTCTTTCGAGCTCTTACAGGCTGCTCTTCTTCCAATCTGTA 420
Db 565 GACTCTTTAGGTAAATGACTGCTTCTTTCGAGCTCTTACAGGCTGCTCTTCTTCCAATCTGTA 624
QY 421 GCAAACTAATCAAAAGCAGCTGAGCTTCTTAAAGAGATGAAGATTAACCCAGTAGTCCCA 480
Db 625 GCAAACTAATCAAAAGCAGCTGAGCTTCTTAAAGAGATGAAGATTAACCCAGTAGTCCCA 684
QY 481 GGGAAACCGCTGCAATGCTCAATCTTATAGTATCAGACAGATGCTACAGCGACACAG 540
Db 685 GGGAAACCGCTGCAATGCTCAATCTTATAGTATCAGACAGATGCTACAGCGACACAG 744
QY 541 ATAGAGAAAGATGAAATGCGATTTAGGATGCAATTTTTCAGGACAGACGCTAGTGGG 600
Db 745 ATAGAGAAAGATGAAATGCGATTTAGGATGCAATTTTTCAGGACAGACGCTAGTGGG 804
QY 601 GCTGTAGAAATGCTAAATCTAATTAACAGTATAGCAACATAGATTTCAGCTTAAGCAGCA 660
Db 805 GCTGTAGAAATGCTAAATCTAATTAACAGTATAGCAACATAGATTTCAGCTTAAGCAGCA 864
QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTTCCAATT 720
Db 865 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTTCCAATT 924
QY 721 CTTCAAGAAAGCGGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 780
Db 925 CTTCAAGAAAGCGGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 984
QY 781 GCAGATGCTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGACAAAGGA 840
Db 985 GCAGATGCTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGACAAAGGA 1044
QY 841 AGTAGTATGCTAGTATTTCCATGCTGTTAGATGATGCTGAAAAATGACACGCT 900
Db 1045 AGTAGTATGCTAGTATTTCCATGCTGTTAGATGATGCTGAAAAATGACACGCT 1104
QY 901 TCCATTTTGTATGCTGGTCTGTCAGATGATTCATGTTCAATAGCGAAATTCCTGAT 960
Db 1105 TCCATTTTGTATGCTGGTCTGTCAGATGATTCATGTTCAATAGCGAAATTCCTGAT 1164
QY 961 TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACGAGCAAGCGCTGGAGAT 1020
Db 1165 TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACGAGCAAGCGCTGGAGAT 1224
QY 1021 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1080
Db 1225 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1284
QY 1081 GCTGGGCAACAAAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG 1140
Db 1285 GCTGGGCAACAAAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG 1344
QY 1141 AGCGCAGAGTTCTCCCGCTGACAGCAAGTTCTATAGGCTCATCTGTAACAGCTTTAC 1200
Db 1345 AGCGCAGAGTTCTCCCGCTGACAGCAAGTTCTATAGGCTCATCTGTAACAGCTTTAC 1404
QY 1201 AAGACTCAAAATCTACAGTTCTGATTTATAAACAACAGATATCAGCAGTTATGATGCT 1260
Db 1405 AAGACTCAAAATCTACAGTTCTGATTTATAAACAACAGATATCAGCAGTTATGATGCT 1464
QY 1261 TAGAAATTCATATGCTATGCTAGGCGAGAAATGATCGGACTCTGATGATGATA 1320
Db 1465 TAGAAATTCATATGCTATGCTAGGCGAGAAATGATCGGACTCTGATGATGATA 1524
QY 1321 AACAACTTAAGTACCCCGCTCTCAGCGATCGCTTCTAGACAGCAACAGAGCTCGA 1380
Db 1525 AACAACTTAAGTACCCCGCTCTCAGCGATCGCTTCTAGACAGCAACAGAGCTCGA 1584
QY 1381 GGACCAAGAAAAACAGATCAAGCGCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTT 1440
Db 1585 GGACCAAGAAAAACAGATCAAGCGCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTT 1644

QY 1441 GGAGATCTCTATAGTCAAGTTTGGGCACTACAACTCTGTATAGCAGATCACTCAGTCAAT 1500
DB 1645 GGAGATCTCTATAGTCAAGTTTGGGCACTACAACTCTGTATAGCAGATCACTCAGTCAAT 1704
QY 1501 CCTCAAGCAATATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCCCTCCA 1560
DB 1705 CCTCAAGCAATATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCCCTCCA 1764
QY 1561 CAGTTTGGCTATCCTTAATGTGCAACTTCTTAATGACTCTACACAGAAATTCATAGCTAAA 1620
DB 1765 CAGTTTGGCTATCCTTAATGTGCAACTTCTTAATGACTCTACACAGAAATTCATAGCTAAA 1824
QY 1621 TTAGAAAGTTTGTGCTGAAGATCTAGACAGCAGCTGAAATTAAGAAAGCACTTCCCTTT 1680
DB 1825 TTAGAAAGTTTGTGCTGAAGATCTAGACAGCAGCTGAAATTAAGAAAGCACTTCCCTTT 1884
QY 1681 GAAACGAACCTCTTGTGTTTATTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATTCGGT 1740
DB 1885 GAAACGAACCTCTTGTGTTTATTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATTCGGT 1944
QY 1741 TATCTCCAATAA 1752
DB 1945 TATCTCCAATAA 1956

RESULT 5
US-09-841-260-63
; Sequence 63, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Brika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-63

Query Match 94.4%; Score 1748.8; DB 10; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGSCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
DB 205 ATGAGTCTGSCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264
QY 61 TTGCGACGCTGGACTCAACGACGAGCGACCGACCTTACGCTCTCTCCACCCAGCTTTGAT 120
DB 265 TTGCGACGCTGGACTCAACGACGAGCGACCGACCTTACGCTCTCTCCACCCAGCTTTGAT 324
QY 121 GATTATAAGACTCAACGCGAAGCAGCTTACGATATCTATCTTTACCTCAACATCAGTCT 180
DB 325 GATTATAAGACTCAACGCGAAGCAGCTTACGATATCTATCTTTACCTCAACATCAGTCT 384
QY 181 GACATACAGCTCTCTTGTGTAGCTTCCAGGATGCTGTCACTAATAAAGATACAGCG 240
DB 385 GACATACAGCTCTCTTGTGTAGCTTCCAGGATGCTGTCACTAATAAAGATACAGCG 444
QY 241 GCTACTGATGAGAAACCGCAATTCGCTGGGAGTGGGAAACTAAGAAATGCCGATGCGATT 300
DB 445 GCTACTGATGAGAAACCGCAATTCGCTGGGAGTGGGAAACTAAGAAATGCCGATGCGATT 504
QY 301 AAAGTTGGCGCGCAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAG 360
DB 505 AAAGTTGGCGCGCAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAG 564

QY 361 GACTCTTTAGGTAAATGACTTCTCTCGACCTCTTACAGGCTCTCTTCTCCAAATCTGTA 420
DB 565 GACTCTTTAGGTAAATGACTTCTCTCGACCTCTTACAGGCTCTCTTCTCCAAATCTGTA 624
QY 421 GCAAAACAATAACAAACAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCCAGTAGTCCCA 480
DB 625 GCAAAACAATAACAAACAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCCAGTAGTCCCA 684
QY 481 GGGAAACCGCTGCAATTCGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
DB 685 GGGAAACCGCTGCAATTCGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 744
QY 541 ATAGAAAGATGGAATCGGATTTAGGATGCAATTTTTCAGGACAGAACCGCTAGTGA 600
DB 745 ATAGAAAGATGGAATCGGATTTAGGATGCAATTTTTCAGGACAGAACCGCTAGTGA 804
QY 601 GCTGTAGAAATGCTTAAATCTAATACAGTATAAGCAACATAGATTACGCTTAAAGCAGCA 660
DB 805 GCTGTAGAAATGCTTAAATCTAATACAGTATAAGCAACATAGATTACGCTTAAAGCAGCA 864
QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCCAAAT 720
DB 865 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCCAAAT 924
QY 721 CTTCAAGAAAGCGGAACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAAATATCAAACT 780
DB 925 CTTCAAGAAAGCGGAACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAAATATCAAACT 984
QY 781 CAGATGTTTCTGATGTTCCAAATCCAGAACTACAGTTTGGAGGCTCCCAAGCAACAGGA 840
DB 985 CAGATGTTTCTGATGTTCCAAATCCAGAACTACAGTTTGGAGGCTCCCAAGCAACAGGA 1044
QY 841 AGTACTATTGTTAGTATTCTGCTTCCATGCTGTTAGATGATGCTGAAAAATCAGACCGCT 900
DB 1045 AGTACTATTGTTAGTATTCTGCTTCCATGCTGTTAGATGATGCTGAAAAATCAGACCGCT 1104
QY 901 TCAATTTTGTATGCTGGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1105 TCAATTTTGTATGCTGGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1164
QY 961 TCTCAAGCTGCCCCAACAGGAGCTCCGACGACAGCTAGAGCGGAAAGCCGCTGGAGAT 1020
DB 1165 TCTCAAGCTGCCCCAACAGGAGCTCCGACGACAGCTAGAGCGGAAAGCCGCTGGAGAT 1224
QY 1021 GACAGTCTGCTGACGCGCTGCGAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1080
DB 1225 GACAGTCTGCTGACGCGCTGCGAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1284
QY 1081 GCTGGGCAACAAACAGGGGCAATCTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG 1140
DB 1285 GCTGGGCAACAAACAGGGGCAATCTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG 1344
QY 1141 AGCGCAGGAGTTCTCCCGCTGACGCAAGTTCTATAGGCTCATCTGTAACAGCTTTAC 1200
DB 1345 AGCGCAGGAGTTCTCCCGCTGACGCAAGTTCTATAGGCTCATCTGTAACAGCTTTAC 1404
QY 1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGAGGTTATGATGCT 1260
DB 1405 AAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGAGGTTATGATGCT 1464
QY 1261 TACAAATCCATCAATGATGCTTGGTAGGGAACGAAATGATGCGACTCGTGTGATGATA 1320
DB 1465 TACAAATCCATCAATGATGCTTGGTAGGGAACGAAATGATGCGACTCGTGTGATGATA 1524
QY 1321 AACATGTAAGTACCCCGCTCTCAGCAGATCCGTTCTCTAGGACGAAACAGAGCTCGA 1380
DB 1525 AACATGTAAGTACCCCGCTCTCAGCAGATCCGTTCTCTAGGACGAAACAGAGCTCGA 1584
QY 1381 GGACCAAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGAACTCTT 1440
DB 1585 GGACCAAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGAACTCTT 1644

QY 1441 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCAATCACTCAGTCGAAT 1500
Db 1645 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCAATCATCCAGTCGAAT 1704
QY 1501 CCTCAGCGAATAATGAGGAGTACAGCAAGCTTACATCGGAGTACAGCAAGCTTCCA 1560
Db 1705 CCTCAGCGAATAATGAGGAGTACAGCAAGCTTACATCGGAGTACAGCAAGCTTCCA 1764
QY 1561 CAGTTTGGCTATCCCTTATGTCGCACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 1620
Db 1765 CAGTTTGGCTATCCCTTATGTCGCACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 1824
QY 1621 TTAGAAAGTTTGTGCTGAGGAGTACAGCAAGCTTACATCGGAGTACAGCAAGCTTCCA 1680
Db 1825 TTAGAAAGTTTGTGCTGAGGAGTACAGCAAGCTTACATCGGAGTACAGCAAGCTTCCA 1884
QY 1681 GAAACGAACTCTCTGTTTATTCAGCAGTGTGCTGCTCAATATCGGCTCTCTATATCTGGT 1740
Db 1885 GAAACGAACTCTCTGTTTATTCAGCAGTGTGCTGCTCAATATCGGCTCTCTATATCTGGT 1944
QY 1741 TATCTCCAAATA 1752
Db 1945 TATCTCCAAATA 1956

RESULT 6

US-10-007-693-63

; Sequence 63, Application US/10007693

; Publication No. US20020146776A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT

; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.515C2

; CURRENT APPLICATION NUMBER: US/10/007,693

; CURRENT FILING DATE: 2001-12-05

; NUMBER OF SEQ ID NOS: 157

; SEQ ID NO 63

; LENGTH: 1956

; TYPE: DNA

; ORGANISM: Chlamydia pneumoniae

US-10-007-693-63

Query Match 94.4%; Score 1748.8; DB 13; Length 1956;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 60
Db 205 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 264
QY 61 TCTGAGAGTGGTCAACGACAGCGACCGCAGCTACGCTCTCCACCCACGTTTGTAT 120
Db 265 TCTGAGAGTGGTCAACGACAGCGACCGCAGCTACGCTCTCCACCCACGTTTGTAT 324
QY 121 GATTATAGACTCAACGCGAACAACAGCTTACGATATCTTTACCTCAACATCACTAGT 180
Db 325 GATTATAGACTCAACGCGAACAACAGCTTACGATATCTTTACCTCAACATCACTAGT 384
QY 181 GACATACAGGCTGCTTGTGAGCTTCCAGGATGCTGCTCAATATTAAGGATACAGCG 240
Db 385 GACATACAGGCTGCTTGTGAGCTTCCAGGATGCTGCTCAATATTAAGGATACAGCG 444
QY 241 GCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGCAGTT 300
Db 445 GCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGCAGTT 504
QY 301 AAGTTGGCGCGCAATTAACAGATTAGCGAAATATGCTTCGGATTAACCAAGCGATTCTT 360
Db 505 AAGTTGGCGCGCAATTAACAGATTAGCGAAATATGCTTCGGATTAACCAAGCGATTCTT 564
QY 361 GACTCTTTAGGTAACAGTACTCTTCTCGACCTTACAGGCTGCTTCTCCAAATCTGTA 420

Db 565 GACTCTTTAGGTAACAGTACTCTTCTCGACCTCTTACAGGCTGCTTCTTCCAAATCTGTA 624
QY 421 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATCAACCCAGTAGTCCCA 480
Db 625 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATCAACCCAGTAGTCCCA 684
QY 481 GGGAAAAACGCTCGCAATGCTCAATCTTATAGTTGATCAGACATGCTACAGGACACAG 540
Db 685 GGGAAAAACGCTCGCAATGCTCAATCTTATAGTTGATCAGACATGCTACAGGACACAG 744
QY 541 ATAGAGAAAGATGGAATGCGATTAGGATGCAATTTTTCAGGACAGCAAGCTTAGTGGGA 600
Db 745 ATAGAGAAAGATGGAATGCGATTAGGATGCAATTTTTCAGGACAGCAAGCTTAGTGGGA 804
QY 601 GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCCAGCTAAAGCAGCA 660
Db 805 GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCCAGCTAAAGCAGCA 864
QY 661 ATGCTTACTGCTAAGACACAAAATAGCTGAAAGCTCAGAAAAAGTTCCCGGACTCTCCAAAT 720
Db 865 ATGCTTACTGCTAAGACACAAAATAGCTGAAAGCTCAGAAAAAGTTCCCGGACTCTCCAAAT 924
QY 721 CTTCAAGAGCGGAACAAATGGTAATACAGGCTGAGAAAGATCTTAAATAATATCAAACT 780
Db 925 CTTCAAGAGCGGAACAAATGGTAATACAGGCTGAGAAAGATCTTAAATAATATCAAACT 984
QY 781 GCAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGCTCCAAAGCAACAAGGA 840
Db 985 GCAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGCTCCAAAGCAACAAGGA 1044
QY 841 AGTAGTATGCTGATGTTCCGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT 900
Db 1045 AGTAGTATGCTGATGTTCCGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT 1104
QY 901 TCCATTTTGTATGCTGCTGGTTTTCGTCAGATGATTCACATGTTCAATACGGAATACTCGAT 960
Db 1105 TCCATTTTGTATGCTGCTGGTTTTCGTCAGATGATTCACATGTTCAATACGGAATACTCGAT 1164
QY 961 TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGAGACGCAAGGCGCTGGAGAT 1020
Db 1165 TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGAGACGCAAGGCGCTGGAGAT 1224
QY 1021 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTTAGGTAAA 1080
Db 1225 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTTAGGTAAA 1284
QY 1081 GCTGGGCAACAAACAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
Db 1285 GCTGGGCAACAAACAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1344
QY 1141 AGCGCAGAGTTCCTCCGCTGAGCAAGTTCTATAGGTCATCTGTAAACACGCTTTAC 1200
Db 1345 AGCGCAGAGTTCCTCCGCTGAGCAAGTTCTATAGGTCATCTGTAAACACGCTTTAC 1404
QY 1201 AAGACCTCAAAATCTACAGGTTCTGATTATTAACACAGATATCAGAGGTTATGATGCT 1260
Db 1405 AAGACCTCAAAATCTACAGGTTCTGATTATTAACACAGATATCAGAGGTTATGATGCT 1464
QY 1261 TACAAATCCATCAATGATGCTTATGAGGCGCAAGAAATGATGCGACTCGTGTATGATA 1320
Db 1465 TACAAATCCATCAATGATGCTTATGAGGCGCAAGAAATGATGCGACTCGTGTATGATA 1524
QY 1321 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGACGCAACAGAGCTCGA 1380
Db 1525 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGACGCAACAGAGCTCGA 1584
QY 1381 GGACAGAAAAAAGATCAGATCAGCCCTCGTAGGGTGTATTTCTGGCAATAGCAGAACTCTT 1440
Db 1585 GGACAGAAAAAAGATCAGATCAGCCCTCGTAGGGTGTATTTCTGGCAATAGCAGAACTCTT 1644
QY 1441 GGAGATGCTATAGTCAAGTTTCGGCACTCAAAATCTGTAATGCAAGTCACTCAGTCGAAT 1500

Db 1645 GGAGATGCTATAGTCAAGTTTCGGCACTCAATCTGTAATGCAGATCATCCAGTCAAT 1704
 QY 1501 CCTCAAGCGAATATAGAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGCCTCCA 1560
 Db 1705 CCTCAAGCGAATATAGAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGCCTCCA 1764
 QY 1561 CAGTTTGGCTATCCTTATGTGCACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 1620
 Db 1765 CAGTTTGGCTATCCTTATGTGCACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 1824
 QY 1621 TTAGAAAGTTTGTGTGCTGAAGGATCTAGGACAGCAGCTGAAATATAAAGCACTTTTCCTTT 1680
 Db 1825 TTAGAAAGTTTGTGTGCTGAAGGATCTAGGACAGCAGCTGAAATATAAAGCACTTTTCCTTT 1884
 QY 1681 GAAACGAACTCCTTGTGTTTATCAGCAGGTGCTGCTGAATATCGGCTCTCTATATTCGGT 1740
 Db 1885 GAAACGAACTCCTTGTGTTTATCAGCAGGTGCTGCTGAATATCGGCTCTCTATATTCGGT 1844
 QY 1741 TATCTCCAAATAA 1752
 Db 1945 TATCTCCAAATAA 1956

RESULT 7

US-10-312-273-14
 ; Sequence 14, Application US/10312273
 ; Publication No. US20040005667A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON Spa
 ; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
 ; FILE REFERENCE: P025035W0
 ; CURRENT APPLICATION NUMBER: US/10/312-273
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: 0616363-4
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 0017047.2
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 0017983.8
 ; PRIOR FILING DATE: 2000-07-21
 ; PRIOR APPLICATION NUMBER: 0019368.0
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: 0020440.4
 ; PRIOR FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: 0022583.9
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 0027549.5
 ; PRIOR FILING DATE: 2000-11-10
 ; PRIOR APPLICATION NUMBER: 0031706.5
 ; NUMBER OF SEQ ID NOS: 664
 ; SOFTWARE: SeqWin99, version 1.02
 ; SEQ ID NO 14
 ; LENGTH: 1956
 ; TYPE: DNA
 ; ORGANISM: Chlamydia pneumoniae

US-10-312-273-14

Query Match 94.4%; Score 1748.8; DB 15; Length 1956;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATTTGCTTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
 Db 205 ATGAGTCTGGCAGATAAGCTGGGTATTTGCTTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264
 QY 61 TCTGCAGAGCTGGACTCAACGACAGCGACCGCACTTACGCTCTCCACCCACGTTTGTAT 120
 Db 265 TCTGCAGAGCTGGACTCAACGACAGCGACCGCACTTACGCTCTCCACCCACGTTTGTAT 324
 QY 121 GATTATAGACTCAACGCGAAGAGCTTAGGATATCTTTACCTTCAACATCACTAGCT 190
 Db 325 GATTATAGACTCAACGCGAAGAGCTTAGGATATCTTTACCTTCAACATCACTAGCT 384

QY 181 GACATACAGGCTGCTTTGGTGGAGCCTCCAGGATGCTGTCTACTAATATAAAGGATACAGCG 240
 Db 385 GACATACAGGCTGCTTTGGTGGAGCCTCCAGGATGCTGTCTACTAATATAAAGGATACAGCG 444
 QY 241 GCTACTGATGAGGAAACCCCAATCGCTGGGAGTGGGAACTAAGAAATGCCGATGCAAGTT 300
 Db 445 GCTACTGATGAGGAAACCCCAATCGCTGGGAGTGGGAACTAAGAAATGCCGATGCAAGTT 504
 QY 301 AAAGTTGGGCGCGCAAAATTACAGAAATATAGCGAAATATGCTTCGGATTAACCAAGCGATTCCT 360
 Db 505 AAAGTTGGGCGCGCAAAATTACAGAAATATGCTTCGGATTAACCAAGCGATTCCT 564
 QY 361 GACTCTTTAGGTAACTGACTTCCCTTCGACCTCTTACAGGCTGCTCTTCTTCCAATCTGTA 420
 Db 565 GACTCTTTAGGTAACTGACTTCCCTTCGACCTCTTACAGGCTGCTCTTCTTCCAATCTGTA 624
 QY 421 GCAAAACAATAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGTCCCA 480
 Db 625 GCAAAACAATAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGTCCCA 684
 QY 481 GGGAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGGACACAG 540
 Db 685 GGGAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGGACACAG 744
 QY 541 ATAGAGAAAGATGGAATCGGATTAGGGATGCAATATTTTGCAGGACAGAACGCTAGTGA 600
 Db 745 ATAGAGAAAGATGGAATCGGATTAGGGATGCAATATTTTGCAGGACAGAACGCTAGTGA 804
 QY 601 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTAAAGCAGCA 660
 Db 805 GCTGTAGAAAATGCTAATACTAATAACAGTATAAGCAACATAGATTTCAGCTAAAGCAGCA 864
 QY 661 ATCGCTACTGCTAAAGACACAAAATAGCTGAAGCTCAGAAAAGTTCCCGACCTCTCCAAAT 720
 Db 865 ATCGCTACTGCTAAAGACACAAAATAGCTGAAGCTCAGAAAAGTTCCCGACCTCTCCAAAT 924
 QY 721 CTTCAAGAGCGGACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 780
 Db 925 CTTCAAGAGCGGACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 984
 QY 781 GCAGATGTTCTGATGTTTCCAAATCCAGGAACTCAGATTGGAGGCTCCAAAGCAACAAGCA 840
 Db 985 GCAGATGTTCTGATGTTTCCAAATCCAGGAACTCAGATTGGAGGCTCCAAAGCAACAAGCA 1044
 QY 841 AGTAGTATTTGGTAGTATTTCGTTTCCATGCTTTTAGATGATGCTGAAAATGAGACCGCT 900
 Db 1045 AGTAGTATTTGGTAGTATTTCGTTTCCATGCTTTTAGATGATGCTGAAAATGAGACCGCT 1104
 QY 901 TCCATTTTGTGCTGGGTTTCGTCAGATGATTACATGTTCAATACGGAATCTCTGAT 960
 Db 1105 TCCATTTTGTGCTGGGTTTCGTCAGATGATTACATGTTCAATACGGAATCTCTGAT 1164
 QY 961 TCTCAAGCTGCCCAACAGGAGCTCCGACGACCAAGCTAGACGACGAAAGCCGCTCGAGAT 1020
 Db 1165 TCTCAAGCTGCCCAACAGGAGCTCCGACGACCAAGCTAGACGACGAAAGCCGCTCGAGAT 1224
 QY 1021 GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTTAGGTAAA 1080
 Db 1225 GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTTAGGTAAA 1284
 QY 1081 GCTGGGCAACACAGGCGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTTTG 1140
 Db 1285 GCTGGGCAACACAGGCGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTTTG 1344
 QY 1141 AGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCTATAGGTCATCTGTAAAAACAGCTTTTAC 1200
 Db 1345 AGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCTATAGGTCATCTGTAAAAACAGCTTTTAC 1404
 QY 1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGAGGTTTATGATGCT 1260
 Db 1405 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGAGGTTTATGATGCT 1464
 QY 1261 TACAAATCCATCATGATGCTCTATGGTAGGGCAAGAAATGATGGACTCTGATGTGATA 1320

Db 1465 TACAAATCCATCAATGATGCCCTATGCTAGGACGAAATGATGCGACTCGTATGATA 1524
Qy 1321 AACAAATGAAATACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAGCTCGA 1380
Db 1525 AACAAATGAAATACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAGCTCGA 1584
Qy 1381 GGACAGAAAAAACAAGATCAAGCCCTCGCTAGGGTGAATTTCTGGCAATAGCAGAACTCTT 1440
Db 1585 GGACAGAAAAAACAAGATCAAGCCCTCGCTAGGGTGAATTTCTGGCAATAGCAGAACTCTT 1644
Qy 1441 GGAGATGCTATAGTCAGATTTCCGACCTACATCTGTAATGACATCACTCAGTCGAAT 1500
Db 1645 GGAGATGCTATAGTCAGATTTCCGACCTACATCTGTAATGACATCACTCAGTCGAAT 1704
Qy 1501 CCTCAAGCGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTCCA 1560
Db 1705 CCTCAAGCGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTCCA 1764
Qy 1561 CAGTTTGGCTATCCCTATGTCGAACTTTCTAATGACTCTACACAGAAATCATAGCTAAA 1620
Db 1765 CAGTTTGGCTATCCCTATGTCGAACTTTCTAATGACTCTACACAGAAATCATAGCTAAA 1824
Qy 1621 TTAGAAAGTTTCTTTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTCCCTT 1680
Db 1825 TTAGAAAGTTTCTTTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTCCCTT 1884
Qy 1681 GAAACGAACCTCTCTTTATTACAGAGGTGCTGATCAATATCGGCTCTCTATATTCCTGT 1740
Db 1885 GAAACGAACCTCTCTTTATTACAGAGGTGCTGATCAATATCGGCTCTCTATATTCCTGT 1944
Qy 1741 TATCTCCCAATA 1752
Db 1945 TATCTCCCAATA 1956

RESULT 8

US-03-564-479-7
; Sequence 7, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (766)..(2235)
US-09-564-479-7

Query Match 62.7%; Score 1161.8; DB 10; Length 2238;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1174; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 ATGAGTCTGGCAGATAAGCTGGGTATTCCTTCTAGTAACAGCTGCTTCTACTAGCAGA 60
Db 970 ATGAGTCTGGCAGATAAGCTGGGTATTCCTTCTAGTAACAGCTGCTTCTACTAGCAGA 1029

Qy 61 TCTCAGACGCTGGACTCAACGACGACGACCGACCTAGCGCTCTCCACCACCGTTGAT 120
Db 1030 TCTCAGACGCTGGACTCAACGACGACGACCGACCTAGCGCTCTCCACCACCGTTGAT 1089
Qy 121 GATTATAGACTCAAGCGCAAAAGAGCTTACGATCTATCTTTACTCTCAACATCAGCT 180
Db 1090 GATTATAGACTCAAGCGCAAAAGAGCTTACGATCTATCTTTACTCTCAACATCAGCT 1149
Qy 181 GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTCTCACTAATAAAGGATACAGCG 240
Db 1150 GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTCTCACTAATAAAGGATACAGCG 1209
Qy 241 GCTACTGATAGGAAAAACGCAATCGCTGGAGTGGAAAACTAAGAAATGCGATGCGATT 300
Db 1210 GCTACTGATAGGAAAAACGCAATCGCTGGAGTGGAAAACTAAGAAATGCGATGCGATT 1269
Qy 301 AAGCTTGGCGCGCAAAATACAGAAATAGCAAAATATGCTTCGGATACCAAGCGATTCTT 360
Db 1270 AAGCTTGGCGCGCAAAATACAGAAATAGCAAAATATGCTTCGGATACCAAGCGATTCTT 1329
Qy 361 GACTCTTTAGTAACTGACTTCTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 420
Db 1330 GACTCTTTAGTAACTGACTTCTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 1389
Qy 421 GCAAACTAATAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Db 1390 GCAAACTAATAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 1449
Qy 481 GGGAAAAAGCTGCAATTTGCTCAATCTTTAGTTCAGACAGATGCTACAGCGACACAG 540
Db 1450 GGGAAAAAGCTGCAATTTGCTCAATCTTTAGTTCAGACAGATGCTACAGCGACACAG 1509
Qy 541 ATAGAGAAAGATGGAATGCAATTTAGGATGCAATTTTTCAGGACAGAAAGCTAGTGA 600
Db 1510 ATAGAGAAAGATGGAATGCAATTTAGGATGCAATTTTTCAGGACAGAAAGCTAGTGA 1569
Qy 601 GCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTGAGCTAAAGCAGCA 660
Db 1570 GCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTGAGCTAAAGCAGCA 1629
Qy 661 ATCGCTACTGCTAAGACACAAATAGCTGAAAGCTCAGAAAAAGTTCCCGACTCTCCAAT 720
Db 1630 ATCGCTACTGCTAAGACACAAATAGCTGAAAGCTCAGAAAAAGTTCCCGACTCTCCAAT 1689
Qy 721 CTTCAAGACGGCAACAAATGGTAATAACAGCTGAGAAAGATCTTAAAAATATCAAACT 780
Db 1690 CTTCAAGACGGCAACAAATGGTAATAACAGCTGAGAAAGATCTTAAAAATATCAAACT 1749
Qy 781 GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTCGAGGCTCCAGCAGCAGGA 840
Db 1750 GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTCGAGGCTCCAGCAGCAGGA 1809
Qy 841 AGTAGTATGCTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGACCGCT 900
Db 1810 AGTAGTATGCTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGACCGCT 1869
Qy 901 TCCATTTTGTGCTGGTTTTCGTCAGATGATTCATGTTCAATGCGAAAAATTCCTGAT 960
Db 1870 TCCATTTTGTGCTGGTTTTCGTCAGATGATTCATGTTCAATGCGAAAAATTCCTGAT 1929
Qy 961 TCTCAAGCTGCCCAACAGGAGCTCGCAGCAGCAGCTAGCAGCAGGAAAGCCGCTGAGAT 1020
Db 1930 TCTCAAGCTGCCCAACAGGAGCTCGCAGCAGCAGCTAGCAGCAGGAAAGCCGCTGAGAT 1989
Qy 1021 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1080
Db 1990 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 2049
Qy 1081 GCTGGGGCAACACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTG 1140
Db 2050 GCTGGGGCAACACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTG 2109
Qy 1141 AGCGCAGGAGTTCTCCCGCTGCGAGCAAGTTCTATAG 1177

Db 2110 AGCGCAGGAG-TACTCCCGCTGACGAAAGTTCTATGG 2145

RESULT 9

US-09-564-479-5
 ; Sequence 5, Application US/09564479
 ; Publication No. US2003095973A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: COHEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; PRIOR FILING DATE: 2000-05-03
 ; PRIOR FILING DATE: 60/132,270
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR FILING DATE: 60/141,276
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 1456
 ; TYPE: DNA
 ; ORGANISM: Chlamydia pneumoniae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(1456)

Query Match 62.1%; Score 1151; DB 10; Length 1456;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 60
 Db 305 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 364
 QY 61 TCTGACAGCTGGACTCAACGACGAGCGGACCTACGCTCCCTCCCTCCACCCAGCTTGTAT 120
 Db 365 TCTGACAGCTGGACTCAACGACGAGCGGACCTACGCTCCCTCCCTCCACCCAGCTTGTAT 424
 QY 121 GATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT 180
 Db 425 GATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT 484
 QY 181 GACATACAGCTGCTTTGGTGAGCTCCAGGATGCTGTCACCTAATAAAGGATACAGCG 240
 Db 485 GACATACAGCTGCTTTGGTGAGCTCCAGGATGCTGTCACCTAATAAAGGATACAGCG 544
 QY 241 GCTACTGATGAGAAACCCCAATCGCTGGGAGTGGAACTAAGAAATCCCGATGCGATT 300
 Db 545 GCTACTGATGAGAAACCCCAATCGCTGGGAGTGGAACTAAGAAATCCCGATGCGATT 604
 QY 301 AAAGTTGGCGCGCAATTAACAGAAATAGCGAAATATGCTTCGATAACCAAGCGATTCTT 360
 Db 605 AAAGTTGGCGCGCAATTAACAGAAATAGCGAAATATGCTTCGATAACCAAGCGATTCTT 664
 QY 361 GACTCTTTAGTAAACTGACTCTTCGACCTCTTACAGCTGCTCTTCTCCAACTGTGA 420
 Db 665 GACTCTTTAGTAAACTGACTCTTCGACCTCTTACAGCTGCTCTTCTCCAACTGTGA 724
 QY 421 GCAACAATAACAAACGACGCTGCTTCTTAAGAGATGCAAGATCAACCGAGTACCCCA 480
 Db 725 GCAACAATAACAAACGACGCTGCTTCTTAAGAGATGCAAGATCAACCGAGTACCCCA 784
 QY 481 GGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCAGACAG 540
 Db 785 GGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCAGACAG 844

QY 541 ATAGAGAAAGATGGAAATGCGATAGGATCATATTTTGCAGGACAGAACGCTAGTGG 600
 Db 845 ATAGAGAAAGATGGAAATGCGATAGGATCATATTTTGCAGGACAGAACGCTAGTGG 904
 QY 601 GCTGTAGAAATGCTTAATCTAATAACAGTATACGACATAGATTCTAGCTTAAGCAGCA 660
 Db 905 GCTGTAGAAATGCTTAATCTAATAACAGTATACGACATAGATTCTAGCTTAAGCAGCA 964
 QY 661 ATCGCTACTGTAAGACACAAATAGCTCAAGCTCAGAAAAAGTTCCCGCACTCTCCAATT 720
 Db 965 ATCGCTACTGTAAGACACAAATAGCTCAAGCTCAGAAAAAGTTCCCGCACTCTCCAATT 1024
 QY 721 CTTCAAGAGCGGAAACAAATGGTAAATACAGCTGAGAAAGATCTTTAAAAATATCAAACT 780
 Db 1025 CTTCAAGAGCGGAAACAAATGGTAAATACAGCTGAGAAAGATCTTTAAAAATATCAAACT 1084
 QY 781 GCAGATGCTTCTGATGTTCCAAATCCAGGAATACAGTTGAGGCTCCAAAGCAAGCA 840
 Db 1085 GCAGATGCTTCTGATGTTCCAAATCCAGGAATACAGTTGAGGCTCCAAAGCAAGCA 1144
 QY 841 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGACCGCT 900
 Db 1145 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGACCGCT 1204
 QY 901 TCCATTTTGTGCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATAATCTGAT 960
 Db 1205 TCCATTTTGTGCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATAATCTGAT 1264
 QY 961 TCTCAAGCTGCCACACAGGAGCTCGCAGCACAAAGCTAGCAGGAAAGCCGCTGGAGAT 1020
 Db 1285 TCTCAAGCTGCCACACAGGAGCTCGCAGCACAAAGCTAGCAGGAAAGCCGCTGGAGAT 1324
 QY 1021 GACAGTCTGCTGACGAGCTGCGAGATGCTCAGAAAGCTTTTAGAAGCGGCTTAGGTAAA 1080
 Db 1375 GACAGTCTGCTGACGAGCTGCGAGATGCTCAGAAAGCTTTTAGAAGCGGCTTAGGTAAA 1384
 QY 1081 GCTGGGCAACAAACAGGCGATCTCAATGCTTTAGACAGATGCTTCTGCTGCTTTGTG 1140
 Db 1385 GCTGGGCAACAAACAGGCGATCTCAATGCTTTAGACAGATGCTTCTGCTGCTTTGTG 1444
 QY 1141 AGCGCAGGAGT 1151
 Db 1445 AGCGCAGGAGT 1455

RESULT 10
 US-09-841-132-57
 ; Sequence 57, Application US/09841132
 ; Patent No. US20020061848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.469C8
 ; CURRENT APPLICATION NUMBER: US/09/841,132
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 599
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 57
 ; LENGTH: 1537
 ; TYPE: DNA
 ; ORGANISM: Chlamydia trachomatis
 US-09-841-132-57

Query Match 10.1%; Score 187.2; DB 9; Length 1537;
 Best Local Similarity 53.0%; Pred. No. 1.7e-42;
 Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;
 QY 863 TTTCCATCTGTTAGATGATGCTGAAAAATGAGACCGCTTCCATTTTGTATGCTGGTTTC 922
 Db 186 TTTCTTTGTTGCTTGTATGATGATGAATGAATGGCAGCGATTCATTCAGGTTTTC 245

Qy	1175	TAGGGTCATCTGTGTAATAACAGCTTTACAGACCTCAAAATCTACAGGTTCTGA-----TT	1228
Db	1361	TCCAGATGAATGTAAACAGCTTTACAGACAGCGGTTCTCTCGACTTCTTCAGGTCCTT	1420
Qy	1229	ATAAAACACAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCCTATGGTA	1288
Db	1421	ATGAGCAGCACTTTCCGATGGATATTCTGCTTCAAAAACACTGAACCTTTATATTCGG	1480
Qy	1289	GGGCACGAATATGATCGCACTCGTGATGTGATPAAACAAATGAAGTACCCCGCTCTCACAC	1348
Db	1481	AAAGCAGAAGCGCGTGCAGTCACCTATTAGTCAACTGCAATCCGCGCTTTCAGAA	1540
Qy	1349	GATCGGTTCTAGAGCAGAAACAGAGCTCGAGGACGAGAAAAACAGATCAAGCCCTCG	1408
Db	1541	CGCTTCTCGTTCTGGCATAGAAAGTCAAGGACGAGTCAGATGCTAGCAAGAGCAG	1600
Qy	1409	CTAGGGTGATTTGGCAATPAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTCGGCAC	1468
Db	1601	CAGAAACTAATGTGCAGAGATAGCCAAACGTTAGGTGATGTATATAGCCGTTACAGGTTTC	1660
Qy	1469	TACAATCTGTAATCGAGATCACTCAGTCGAATCCTCAAGCGAATAATGAGGAGATCAGAC	1528
Db	1661	TGGATTCTTTGATGCTACGATTTGTGAGCAATCCGCAAGTAATCAGAAGAGATTATGC	1720
Qy	1529	AAAAGCTTACATCGGCAGTGACAAAGCCTCCACAGTTTGGCTATCTCTTATGTGCCAATTT	1588
Db	1721	AGAAGCTCAGGCATCTATTAGCAAGCTCCACAAATTTGGGTATCTCTGCTGTTCAGAAAT	1780
Qy	1589	CTAATGACTTACACAGAGTTCACTAGCTAAATTAGAAAGTTTGTGTTCTGTAAGGATCTA	1648
Db	1781	CTCGGGATAGCTTGCAGAAAGTTTGCTCGCAATTGGAAAGAGAGTTTGTGTATGGGGAAC	1840
Qy	1649	GGACAGCAGCTGAAATAAAGCACTTTCCTTTGAAACGCACTCCTTGTTTATTTCAGCAGG	1708
Db	1841	GTAGTCTCGAGAACTCTCGAGAGATCGGTTTAGAAAAACGCCCGCTTCAITTCACAGG	1900
Qy	1709	TGCTGGTCAATATCGGCTCTCTATATTCTGGTTATCT	1745
Db	1901	TGTTGGTAAACATGTCTTCTCTATCTTCGTGTTATCT	1937

RESULT 12

```

US-09-841-260-13
; Sequence 13, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Brika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121-515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 13
; LENGTH: 1171
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-841-260-13

```

	Query Match 9.9%; Score 184; DB 10; Length 1171; Best Local Similarity 52.8%; Pred. No. 1.2e-41; Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;
QY	863 TTTCACGCTGTTAGATGATGCTGAAATAGACCGCTTCCATTTTGATGTTCTGGGTTTC 922
Db	 207 TTTCCTTGTTGCTTGATGATGTAGACAATGAAATGGCAGCAATGCACCTGCAAGGTTTTTC 266
QY	923 GTCAGATGATTCATCANGTTTCAATACGGAAATCCTGATTTCTACAGCTGCCCAACGAGAGC 982
Db	267 GATCTGATGTCGAACAATTTAATGTAACAACATCCTGCAACAGCTAAAGAGAGCTACAAGCTA 326

RESULT 13

US-10-007-693-13
 ; Sequence 13, Application US/10007693
 ; Publication No. US20020146776A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
 ; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.515C2
 ; CURRENT APPLICATION NUMBER: US/10/007,693
 ; CURRENT FILING DATE: 2001-12-05
 ; NUMBER OF SEQ ID NOS: 157
 ; SEQ ID NO 13
 ; LENGTH: 1171

```
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-10-007-693-13

Query Match          9.9%; Score 184; DB 13; Length 1171;
Best Local Similarity 52.8%; Pred. No. 1.2e-41;
Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;

Qy 863 TTTCCATGCTGTAGATGATGCTGAAATGAGACCGCTTCCATTTTGATGCTCGGTTTC 922
Db      |||||
Qy 207 TTTCTTGTCTTGATGATGACATGAAATGCGACGATTCGATCGCAAGTTTC 266
Db      |||||
Qy 923 GTCAGATGATTCACATGTTCAATACGGAATTCCTGATTCCTCAAGCTCCCAACAGGAGC 982
Db      |||||
Qy 267 GATCTATGATCGAACAATTAATGTAACAATCTCTGCAACAGCTAAAGAGCTACAAGCTA 326
Db      |||||
Qy 983 TCGCAGCAACAAGCTAGACGACGACGCGCTGGAGATGACAGTCTCTCGACGCGTG 1042
Db      |||||
Qy 327 TGGAGGCTCAGCTGA---CTGCGATGTCAGATCAACTCTGTTGGTGGGATGGCGAGCTCC 383
Db      |||||
Qy 1043 CAGATGCTCAGAAAGCTTTAGAAGCGGCTTAGGTAAAGCTGGGCAACAAC----- 1093
Db      |||||
Qy 384 CAGCGAAATACAAGCAATCAAGATGCTCTTGGCAGACTTTGAAACAACCATCAGCAG 443
Db      |||||
Qy 1094 AGGGCACTACTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTTGTTGAGCGCAGAGTTC 1153
Db      |||||
Qy 444 ATGGTTTGGCTACAGCTATGGACAAGTGGCTTTTTCAGCTGCCAAGGTTGGAGGAGCT 503
Db      |||||
Qy 1154 CTCGCGCTGACGACGATGTTCTATAGGTCATCTGTAAACAGCTTTTACAAAGCTCAAAAT 1213
Db      |||||
Qy 504 CCGCAGCAACAGCTGCGCACTCTCCAGATGAATGTAAACAGCTTTTACAGACAGGTTT 563
Db      |||||
Qy 1214 CTACAGGTTCT-----GATTATAAACAACAGATATCAGAGGTTATGATGCTTACAAAT 1267
Db      |||||
Qy 564 CTTCAGCTTCTTCCAGCTCTTATGACGACGACATTTCCGATGGATATCTGCTTACAAA 623
Db      |||||
Qy 1268 CCATCAATGATGCCCTATGTTAGGACGAAATGATGCGACTCGTGATGTATAACAATG 1327
Db      |||||
Qy 624 CACTGAATCTCTTTATTCGAAAGCAGAAAGCGGCTGACGCTGATTAAGTCAAACTG 683
Db      |||||
Qy 1328 TAAGTACCCCGCTCTCACAGATCCGTTTCTTAGACGACGAAAGCTCGAGGACCG 1387
Db      |||||
Qy 684 CAATCCCGCTTTCAGAGAGCTTTCTCGTTCTGGCATAGAAAGTCAAGGACGAGTG 743
Db      |||||
Qy 1388 AAAAACAAGATCAAGCCCTCGTAGGGTGAATTTCTGGCAATAGCAGAACTCTTGAGATG 1447
Db      |||||
Qy 744 CAGATGCTTAGCCAAAGACGACGACGAAACTATTTGTCAGAGATAGCCAAACGTTAGTGTATG 803
Db      |||||
Qy 1448 TCTATAGTCAAGTTTTCGGCACTTACAACTCTGTAATGACAGTCACTCAGTCTCGAATCCTCAAG 1507
Db      |||||
Qy 804 TATATAGCGCTTACAGTTCTTGGATTTCTTTGATGCTACGATTTGTAGCAATCCGCAAG 863
Db      |||||
Qy 1508 CGAATATAGGAGATCAGCAAAAGCTTACATCGGCACTGACAAAGCTCCACAGTTTG 1567
Db      |||||
Qy 864 CAATCAAGAAAGAGATTAATGACAGAGCTCAGGCACTATTTAGCAAAAGCTCCACAAATTG 923
Db      |||||
Qy 1568 GCTATCTTATGTGCACTTCTTAATGACTCTACACAGAAAGTTCAATGATTAATAGAAA 1627
Db      |||||
Qy 924 GGTATCTCTGCTTTCAGAAATCTGCGGATAGCTTTGACAGAGTTTGTCTCGCAATTTGAAA 983
Db      |||||
Qy 1628 GTTTCTTTTGTGTAAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTCTTTTGAACGA 1687
Db      |||||
Qy 984 GAGATTTTGTGATGGGAACGTAGTCTCGCAGAAATCTCAAGAGAAATCGTTTATAGAAAC 1043
Db      |||||
Qy 1688 ACTCCTTGTATTCAGCAGGTTGCTGGTCAATATCGGCTCTCTAATATCTGTTATCTTCC 1747
Db      |||||
Qy 1044 AGCCCGCTTTCATTCAACAGGTTTGGTAAACATTTGCTTCTATCTCTGTTATCTTT 1103
Db      |||||
Qy 1748 AATAAC 1753
Db      |||||
Qy 1104 CTTAAC 1109
Db      |||||
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:19 ; Search time 4855.74 Seconds

(without alignments)

11389.568 Million cell updates/sec

Title: US-10-608-559-3

Perfect score: 1852

Sequence: 1 atgagtgcgcagataagct.....gatcttcaataataaaaga 1852

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estnu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_est3.*

12: gb_est4.*

13: gb_est5.*

14: gb_est6.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	52.8	2.9	843	28	BZ643413	BZ643413. OGAOW57TM
c 2	48.4	2.6	1191	14	CD500670	CD500670. GDA46-D01
c 3	48	2.6	872	14	CK097864	CK097864. UB64CFC02
c 4	46.6	2.5	359	9	AL840303	AL840303. AL840303

c 5	46.2	2.5	598	14	CD661508	CD661508
c 6	45.6	2.5	1201	13	EX381961	EX381961
c 7	45.4	2.5	528	12	BJ340195	BJ340195
c 8	45.4	2.5	724	12	BJ344100	BJ344100
c 9	45.2	2.4	336	14	CB098017	CB098017
c 10	45.2	2.4	817	14	CD656959	CD656959
c 11	45.2	2.4	856	13	BUI75987	BUI75987
c 12	45.2	2.4	890	13	BQ424404	BQ424404
c 13	44.8	2.4	484	9	AA314486	AA314486
c 14	44.8	2.4	885	13	EX425603	EX425603
c 15	44.4	2.4	856	28	BH132709	BH132709
c 16	44	2.4	1869	29	AY406745	AY406745
c 17	43.8	2.4	745	12	BJ430078	BJ430078
c 18	43.6	2.4	701	12	BJ372526	BJ372526
c 19	43.6	2.4	735	12	BJ398404	BJ398404
c 20	43.4	2.3	443	29	FR0008252	FR0008252
c 21	43.4	2.3	801	28	BH314010	BH314010
c 22	43.4	2.3	850	9	AV406016	AV406016
c 23	43.4	2.3	861	28	AZ682538	AZ682538
c 24	43.4	2.3	896	12	BP126288	BP126288
c 25	43.4	2.3	906	28	AZ550193	AZ550193
c 26	43.4	2.3	963	12	BP126164	BP126164
c 27	43.2	2.3	404	14	CF587772	CF587772
c 28	43	2.3	420	14	CB798245	CB798245
c 29	43	2.3	578	13	BU709558	BU709558
c 30	43	2.3	711	14	CB055699	CB055699
c 31	43	2.3	756	14	CA326603	CA326603
c 32	43	2.3	825	14	CF748551	CF748551
c 33	43	2.3	931	13	BQ000093	BQ000093
c 34	43	2.3	938	13	BU559883	BU559883
c 35	42.8	2.3	606	12	BJ314050	BJ314050
c 36	42.8	2.3	616	28	AZ525704	AZ525704
c 37	42.8	2.3	764	28	CC017401	CC017401
c 38	42.8	2.3	1124	13	BX436282	BX436282
c 39	42.6	2.3	473	12	BI896076	BI896076
c 40	42.6	2.3	536	14	CD569533	CD569533
c 41	42.6	2.3	635	13	BW056885	BW056885
c 42	42.4	2.3	656	12	BQ10829	BQ10829
c 43	42.4	2.3	701	12	BJ490326	BJ490326
c 44	42.4	2.3	980	29	CNS01PHD	CNS01PHD
c 45	42.2	2.3	843	29	CC569909	CC569909

ALIGNMENTS

RESULT 1
BZ643413/c
LOCUS BZ643413 843 bp DNA linear GSS 29-JAN-2003
DEFINITION OGAOW57TM_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBMA011J17,
genomic survey sequence.
ACCESSION BZ643413
VERSION BZ643413.1 GI:28104915
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 843)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGAOW57C
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR

[illegible]

FEATURES
 source
 Email: bo.segerman@plantphys.umu.se.
 Location/Qualifiers
 1..872

/organism="Populus tremula"
 /mol_type="mRNA"
 /db_xref="taxon:113636"
 /clone="UB64CPC02"
 /tissue_type="Active cambium"
 /clone_lib="Populus active cambium cDNA library"

ORIGIN

Query Match 2.6%; Score 48; DB 14; Length 872;
 Best Local Similarity 50.4%; Pred. No. 0.11; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
 QY 957 TGATTTCAAGCTGCCCAACAGGAGCTCGCAGCACAAAGCTAGAGCAGCAGCAAGCGCTGG 1016
 Db 513 TGTGCTGCTGCTGCTCAATTGTTGGAGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGC 572
 QY 1017 AGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTTAGG 1076
 Db 573 TGTGCAATTGAGGTAGTCTGCTGTGGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGCTGC 632
 QY 1077 TAAAGCTGGGCAACAAAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1136
 Db 633 TGTGGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGCTGCTGCTGCTGCTGTTGAAGC 692
 QY 1137 TGTGAGCGCAGGAGTCTCTCCCGCTGCAGCAAGTTCTATAGGGTCATCTGTA 1188
 Db 693 TGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGTTGTA 744

RESULT 4
 AL840303
 LOCUS
 DEFINITION AL840303 F000C Takifugu rubripes cDNA clone F000C10aC11, mRNA
 sequence.
 ACCESSION AL840303.1 GI:21892241
 VERSION AL840303.1
 KEYWORDS EST.
 SOURCE Takifugu rubripes (Fugu rubripes)

ORGANISM
 Takifugu rubripes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Takifugu.
 1 (bases 1 to 359)
 Clark.M.S.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Unpublished (2002)
 Contact: Clark MS
 MRC Human Genome Mapping Project Resource Centre
 Hinxton, Cambridge, CB10 1SB, UK
 Email: biohelp@hgmpr.mrc.ac.uk
 Vector: pME18S-FL3
 V_type: phagemid
 PRIMER: ME-735FW
 Library created by Koichi Kawakami, Masahide Sasaki, Yutaka Suzuki,
 Sumio Sugano
 The Institute of Medical Science, The University of Tokyo,
 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
 and

Kiyoshi Kikuchi, Shugo Watabe
 Laboratory of Aquatic Molecular Biology and Biotechnology, Graduate
 School of Agricultural and Life Sciences, The University of Tokyo,
 Bunkyo-ku, Tokyo 108-8639, Japan
 Library sequenced by Melody S. Clark and Amanda Thompson MRC Human
 Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10
 1SB, UK.

FEATURES

source
 1..359
 /organism="Takifugu rubripes"
 /mol_type="mRNA"

/db_xref="taxon:31033"
 /clone="F000C10aC11"
 /sex="female"
 /tissue_type="ovary"
 /dev_stage="adult"
 /clone_lib="F000C"
 /note="Vector: pME18S-FL3"

ORIGIN

Query Match 2.5%; Score 46.6; DB 9; Length 359;
 Best Local Similarity 49.8%; Pred. No. 0.16;
 Matches 118; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
 QY 970 GCCCAACAGAGCTCCAGCACAAAGCTAGAGCAGCAAGCGCTGAGATGACAGTGTCT 1029
 Db 47 GCGAAGGAGCAGAGAGCTGCTGAAGCTGCGAGGAGCAAGAGCTGCTGAAGCTGTGAAGCT 106
 QY 1030 GCTGCAGCGCTGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGTTAAAGCTGGCAA 1089
 Db 107 GCAGGAGCAGAGCTGCAGGAGCAGAGCTGCGAGGAGCAGAGCTGCTGAAGCTGCAGGA 166
 QY 1090 CAACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGAGCGCAGGA 1149
 Db 167 GCAGAAAGCTGCAGAGCTGCAGGAGCAGAGCTGCTGAAGCTGCTGAGCAGAGCTGCA 226
 QY 1150 GTTCTCTCCCGCTGCAGCAAGTTCTATAGGCTCATCTGTTAAACAGCTTTTCAAGACC 1206
 Db 227 GGAGCAGAGCTGCTGAAGTGCGGAGCAGAGAGCTGCTGAAGTCTTGAAGGGTCC 283

RESULT 5

CD661508/c
 LOCUS
 DEFINITION CD661508 598 bp mRNA linear EST 19-JUN-2003
 cDNA 5' similar to TR:Q9UPM8 Q9UPM8 EPSILON-ADAPTIN. i, mRNA
 sequence.

ACCESSION CD661508.1 GI:31997248
 VERSION CD661508.1
 KEYWORDS EST.
 SOURCE Eimeria tenella
 ORGANISM Eimeria tenella
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
 Eimeria.

REFERENCE
 AUTHORS
 1 (bases 1 to 598)
 Liberato, P., Diaz, C., Tang, K., Marra, M., Hillier, L., Kucaba, T.,
 Martin, J., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Florence, N., Shin, I., Jackson, Y.,
 Cardenas, M., McCann, R., Waterston, R., Wilson, R. and Sibley, D.

WashU-Merck Eimeria tenella project
 Unpublished (1999)
 Contact: David Sibley, Ph.D.
 WashU-Merck Eimeria tenella project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu

Contact David Sibley (toxoeat@borcim.wustl.edu) for further
 information relating to organism, libraries, or clone availability.
 Seq primer: -40RP from Gibco.

FEATURES

source
 1..598
 /organism="Eimeria tenella"
 /mol_type="mRNA"
 /db_xref="taxon:5802"
 /dev_stage="LS18"
 /lab_host="E.coli DH10B (GeneHog, Invitrogen, Inc)"
 /clone_lib="Eimeria tenella S5-2 Excised cDNA"
 /note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; Sporozoites were obtained from in vitro sporulated
 and excysted oocysts of E. tenella grown in chickens. cDNA

was synthesized from poly mRNA using an oligo-dT primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA and products were size-selected on sephacryl S500. The cDNA were ligated to EcoRI/XhoI prepared lambda ZapII (Stratagene). The primary library was mass excised using Exasist helper phage (Stratagene). The phagemids were precipitated with PEG 8000, extracted with phenol/chloroform and electroporated into DH10B cells. The library may contain a small percentage of host or bacterial contaminants."

ORIGIN

Query Match 2.5%; Score 46.2; DB 14; Length 598;
Best Local Similarity 49.0%; Pred. No. 0.28;
Matches 123; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 916 GGGTTTCGTGAGATTCACATGTTCAATACGGAATCTCTCAAGTCCCA 975
DB 343 GGGTTTAGGGTTATGATGAGATAGGCGCGTGGCTAAACCCCTAGGCTTTATAA 284

QY 976 CAGGAGCTCGCAGCACAACTAGACGAGCGAAGCGCTGGAGATGACAGTGTGTGCA 1035
DB 283 CCCAGGTTTCGGTTGCTGCTGCTCCAGAGGCTTCAGCCGACGACGCCGCT 224

QY 1036 GCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTTAGGTAAGCTGGGCAACACAG 1095
DB 223 GCAGCGGCTGCTGTCGGCAGTTGAAGATGCTGTTCTATTGCTGCTGCTGCTGTC 164

QY 1096 GGCATACCTCAATGCTTAGACAGATGCTGCTGCTGCTGCTGCTGAGCGAGGATTCCT 1155
DB 163 CTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104

QY 1156 CCGCTGCGAGC 1166
DB 103 GCTGCTGCGGC 93

RESULT 6

BX381961 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSQDI072VF05 3-PRIME, mRNA sequence.

ACCESSION BX381961 GI:30453007

VERSION EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSQDI072CC03NP1.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSQDI072VF05"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="First strand cDNA was primed with a NotI-oligo(dn)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

ORIGIN

Query Match 2.5%; Score 45.6; DB 13; Length 1201;
Best Local Similarity 4.3%; Pred. No. 0.62;
Matches 30; Conservative 243; Mismatches 430; Indels 0; Gaps 0;

QY 740 TGGTAATACAGCGCTGAGAAAGATCTTAAATAATCAACCTGCAGATGGTTCGTGATGTC 799

DB 391 KKKKACANNNNNKCKMMNNNNKKKKCMNNKKNNNNKKNNNNKKNNNNKKNNNNNNNNNA 450

QY 800 CAAATCCAGGAACACAGTTCGAGGCTCCAAGCAACAGGAAGTAGTAGTGTGATGATTC 859

DB 451 NMNNNNKKNN 510

QY 860 GTGTTTCCATGCTGTAGATGATGCTGAAATGAGACCGCTTCATTTTGTGATGCTGGGT 919

DB 511 CKMMNNNNKMMNN 570

QY 920 TTCGTGAGATGATTCACATGTTCAATACGGAATCTCTGATTCCTCAAGCTGCCAACAGG 979

DB 571 NN 630

QY 980 AGCTCGCAGCAACAGCTAGACGAGCAAGCGCTGGAGATGACAGTGTGCTGCGAGCGC 1039

DB 631 KNN 690

QY 1040 TGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAGAGTGGGCAACACAGGCA 1099

DB 691 NNN 750

QY 1100 TACTCAATGCTTAGAGCAGATCGCTTCTGCTGCTGTTGTGAGCGCAGAGTTCCTCCG 1159

DB 751 NNNTNN 810

QY 1160 CTGACGCAAGTTCATAGGGTTCATCTGTAACACAGCTTTTACAGACCTCAAAATCTACAG 1219

DB 811 BKGMVCMCKMMNN 870

QY 1220 GTTCTGATTATAAACACAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATG 1279

DB 871 KDNMMNN 930

QY 1280 CCTATGGTAGGCGACGAAATGATGCGACTGCTGTGTGATAAACAATGTAAGTACCCCG 1339

DB 931 KKKMMNN 990

QY 1340 CTCTCACAGATCCGTTCTAGAGCAGCAACAGAGAGCTCGAGACAGAAAAAACAATC 1399

DB 991 MCWCKKKBKMKVCKCKCKMMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1050

QY 1400 AAGCCCTCCCTAGGTTGATTTCTGCAATAGCAGAACTCTTGG 1442

DB 1051 KKKKCKKKBKMMNN 1093

RESULT 7

BX340195/c

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

BX340195 528 bp mRNA linear EST 07-MAR-2002

BX340195 Dictyostelium discoideum cDNA library, AF Dictyostelium

discoideum cDNA clone dda13f05 3', mRNA sequence.

BX340195

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source
1. .528
Location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddal2f05"
/sex="mat A"
/dev_stage="Aggregation stage"
/clone_lib="Dictyostelium discoideum cDNA library, AF"

ORIGIN

Query Match 2.5%; Score 45.4; DB 12; Length 528;
Best Local Similarity 44.8%; Pred. No. 0.43; Indels 0; Gaps 0;
Matches 175; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 514 GATCAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGGATTAGGGATGCA 573
DB 480 GATGCCAAGAAAGTGTGTGTCAGAGCAAGAAAGCTGCAGAAAGAAAGCTGCT 421

QY 574 TATTTTGCAGCAGCAAGCGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATA 633
DB 420 GATGCAGAAAGCAAGAAAGCTGCTGATGCCAAGAAAGCTGCTGCTGATGAAGAAAGCA 361

QY 634 AGCAACATAGATTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 693
DB 360 AAGAAAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 301

QY 694 CAGAAAAAGTTCCCGACTCTCCAAATCTTCAAGAAGCGGAACAAATGGTAATACAGCT 753
DB 300 GCAAGAAAGCTGCTGATGCCAAGAAAGTGTGTGTCAGAGCAAGAAAGCTGCTGAT 241

QY 754 GAGAAAGATCTTAAAAATATCAACCTGCAGATGTTCTGATGTTCCAAATCCAGGAAT 813
DB 240 GCTAAGAAAGCTGCTGATGAAGAAAGCAAGAAAGCTGCTGATGCTGATGCTGATGCTGCT 181

QY 814 ACAGTTGGAGCTCAACACACAGGAAGTAGTATTGGTAGTATTGGTATTTCATGCTG 873
DB 180 GAAGAAAGAAAGCAAGAAAGCTGCTGATATCAAGAAAGCTGCAGAAAGATGCCAAAAA 121

QY 874 TTAGATGATGCTGAAAAATGAGACCGCTTCCA 904
DB 120 GCAGAGATGCCAAAAAGCAGAAAGATGCCA 90

RESULT B
BU344100/c 724 bp mRNA linear EST 06-MAR-2002
LOCUS
DEFINITION
Dictyostelium discoideum cDNA library, AF Dictyostelium
discoideum cDNA clone ddal2f03, mRNA sequence.

ACCESSION
BU344100
VERSION
BU344100.1 GI:19214607
KEYWORDS
EST.

SOURCE
Dictyostelium discoideum

ORGANISM
Dictyostelium discoideum

REFERENCE
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

1 (bases 1 to 724)

Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

Full length cDNA of Dictyostelium discoideum at the aggregation

stage

Unpublished (2002)

CONTACT: Tadashi Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

FEATURES

source
1. .724
Location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddal2f05"
/sex="mat A"
/dev_stage="Aggregation stage"
/clone_lib="Dictyostelium discoideum cDNA library, AF"

ORIGIN

Query Match 2.5%; Score 45.4; DB 12; Length 724;
Best Local Similarity 44.8%; Pred. No. 0.52; Indels 0; Gaps 0;
Matches 175; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 514 GATCAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGGATTAGGGATGCA 573
DB 679 GATGCCAAGAAAGTGTGTGTCAGAGCAAGAAAGCTGCAGAAAGAAAGCTGCT 620

QY 574 TATTTTGCAGCAGCAAGCGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATA 633
DB 619 GATGCAGAAAGCAAGAAAGCTGCTGATGCCAAGAAAGCTGCTGCTGATGAAGAAAGCA 560

QY 634 AGCAACATAGATTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 693
DB 559 AAGAAAGCTGCTGATGCCAAAAAGCTGCCGATGCCAAGAAAGCTGCTGATGAGAAAGAA 500

QY 694 CAGAAAAAGTTCCCGACTCTCCAAATCTTCAAGAAGCGGAACAAATGGTAATACAGCT 753
DB 499 GCAAGAAAGCTGCTGATGCCAAGAAAGTGTGTGTCAGAGCAAGAAAGCTGCTGAT 440

QY 754 GAGAAAGATCTTAAAAATATCAACCTGCAGATGTTCTGATGTTCCAAATCCAGGAAT 813
DB 439 GCTAAGAAAGCTGCTGATGAAGAAAGCAAGAAAGCTGCTGATGCTGATGCTGATGCTGCT 380

QY 814 ACAGTTGGAGCTCCAGCAACAGGAAGTAGTATTGGTAGTATTGGTATTTCATGCTG 873
DB 379 GAAGAAAGAAAGCAAGAAAGCTGCTGATATCAAGAAAGCTGCAGAAAGATGCCAAAAA 320

QY 874 TTAGATGATGCTGAAAAATGAGACCGCTTCCA 904
DB 319 GCAGAGATGCCAAAAAGCAGAAAGATGCCA 289

RESULT 9

CB098017

LOCUS

DEFINITION

Strongyloides ratti cDNA 5' similar to contains element XTR

repetitive element ;, mRNA sequence.

CB098017

VERSION

CB098017.1 GI:27922200

KEYWORDS

SOURCE

ORGANISM

Strongyloides ratti

Strongyloides ratti

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Panagrolaimidae; Strongyloidea; Strongyloides.

1 (bases 1 to 336)

McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,

Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,

Tsagaris, V., R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,

Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,

Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

CONTACT: McCarter, J.P

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810
Email: estowatson.wustl.edu
The library was constructed by Claire Murphy, Brandi Chiapelli and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dynal). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis kit from Clontech. Directionally cloned into the UDG sites of pAMP1. Parasitic adult females were collected from immunologically naive animals and provided by Dr. Mark Viney of Bristol, UK.
Seq primer: -40RP from Gibco
High quality sequence stop: 298.

FEATURES

source

```
1. 336
/organism="Strongyloides ratti"
/mol_type="mRNA"
/db_xref="taxon:34506"
/dev_stage="parasitic adult females"
/lab_host="DH10B"
/clone_lib="Strongyloides ratti PA female naive pAMP1 v1"
/notes="Vector: pAMP1 (Gibco); Site1: NotI; Site2: SalI;
The library was constructed by Claire Murphy, Brandi
Chiapelli and Dr. James McCarter at Washington University,
St. Louis. The cDNA was made by using Dynabead oligo-dT
priming (Dynal). PCR based library using a modified
protocol from the SMART PCR cDNA Synthesis kit from
Clontech. Directionally cloned into the UDG sites of
pAMP1. Parasitic adult females were collected from naive
animals and provided by Dr. Mark Viney of Bristol, UK."
```

ORIGIN

```
Query Match      2.4%; Score 45.2; DB 14; Length 336;
Best Local Similarity 47.8%; Pred. No. 0.38;
Matches 131; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 544 GAGAAAGATGGAATGGGATAGGATGATATTTTCAGACAGACAGCGTAGTGGAGCT 603
Db 31 GAAACAGAGATATCTGTTCTTAAGAAACTGAAGTTACAGAAACTTCTGGTACAGAGAA 90

Qy 604 GTAGAAATGTAATCTAATACAGTATAGCAACATAGATTCAGCTTAAGCAGCAATC 663
Db 91 ACAGAAATGTCGAATCTAAGAAACTGAATTCCTGAATCTTCTGGTACAGAAAGC 150

Qy 664 GCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTTCCCAATTCCT 723
Db 151 GAATTCATATCTGACAAACAGAAACTCTCTTACATATCCAGATATTCAGGTAGC 210

Qy 724 CAAGAAGCGGACAAATGGTAAATACAGCTGAGAAAGATCTTAAANATCAACACCTGCA 783
Db 211 ACTGAATCTAGAGAAACAGAAAGTATCTGCTTCTTAAGAAACTGAAGTCACTGAAACTCT 270

Qy 784 GATGTTCTGATGTTCCAAATCCAGGAAGTACAG 817
Db 271 GCCACAGAGAAACTGAATCTGTAATCTTAAG 304
```

RESULT 10

```
CD656959/c
LOCUS      CD656959      817 bp      mRNA      linear      EST 18-JUN-2003
DEFINITION AGENCOURT_14540498 NIA Human H1 Embryonic Stem Cell cDNA Library
            (Long) Homo sapiens cDNA clone IMAGE:30421464 5', mRNA sequence.
ACCESSION  CD656959
VERSION    CD656959.1 GI:31898704
KEYWORDS  EST.
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 817)
            NIH-MGC http://mgs.nci.nih.gov/
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL   JOURNAL
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
```

National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9pbbs-remail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM499 row: f column: 01
High quality sequence stop: 760.
Location/Qualifiers

FEATURES

source

```
1. 817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30421464"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library
(Long)"
/notes="Vector: pCMV-Sport6; Site1: NotI; Site2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MBF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(invitrogen/Gibco) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with Trizol Reagent
from invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:11544199] Double-stranded cDNAs were
synthesized with an oligo(dT) primer (Invitrogen:
5'-pGACTAGTCTAGATCCGAGCGCGCCCTTTTCTTTT-3') from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker LL-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-SPORT6 plasmid vector. The average insert size is
about 3.6kb."
```

ORIGIN

```
Query Match      2.4%; Score 45.2; DB 14; Length 817;
Best Local Similarity 45.3%; Pred. No. 0.64;
Matches 164; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

Qy 954 TCTGATTCACAGTCCCAACAGAGCTCGCAGCACACAGCTAGACGACGAAAGCCGC 1013
Db 571 TCTGGAATCTCCACTGCTGGATCATGTCCAAAAAGCTGCTGCTGAGCGAGTGGT 512

Qy 1014 TGGAGATGACAGTGTCTGCGCGCTGGCAGATGCTCAGAAAGCTTTAGAACGGCTCT 1073
Db 511 GGCTGCTGTTGCTGCTGCTGCTGCGGGGCTGCTGAGTGGAACTGATGCTCAGGGCC 452

Qy 1074 AGGTAAGACTGGGCACACACAGGGGCATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGC 1133
Db 451 TGTGACCTCTGACTCAAGTGTGGCGCTGTCTGACTTTGCTGCTGCTGCTGCTGCTGC 392

Qy 1134 TGTGTGAGCGCAGAGGAGTCTCCCGCTGCAGCAAGTCTCTATAGGTCATCTCTAAACA 1193
```

```

Db      391  TGGTGCATCGCTGGAGCTGCTGCTGTAATCTGCTGTTGCTGCTGCTGGAGCAGG 332
QY      1194  GCCTTACAGACCTCAAAATCTACAGCTTCTGATATATAAAACACAGATATCAGCAGGTTA 1253
Db      331  CTGTGCTGCTGCTGGAGCTGGCAGCTGCTGCTGCTGGAGCTGGTCTGCTGCAGCTGA 272
QY      1254  TGATGCTTACAATCCATCAATGATGCTATGTTAGGCGACGAAATGATGCGACTCGTGA 1313
Db      271  TGGAACTGCTGCTGCAGGCGATGCTGCTGCTGAAAGGCTGGAGCTGCTGCTGAGGCGGA 212
QY      1314  TG 1315
Db      211  TG 210

RESULT 11
LOCUS   BUI75987              856 bp      mRNA      linear      EST 04-SEP-2002
DEFINITION   AGENCOURT_7782344 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6137044
5', mRNA sequence.
BUI75987
VERSION      BUI75987.1 GI:22689971
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 856)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13450 row: d column: 05
            High quality sequence stop: 602.
FEATURES    Location/Qualifiers
             source
              1..856
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6137044"
               /tissue_type="retinoblastoma"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH_MGC_67"
               /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
               Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
               Average insert size 1.75 kb. Library constructed by Life
               Technologies."
ORIGIN
Query Match      2.4%; Score 45.2; DB 13; Length 856;
Best Local Similarity 45.3%; Pred. No. 0.66;
Matches 164; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY      954  TCGTGAATCTCAAGTCCCAACAGAGCTCGACACACAGTACAGCAGCAAGCGC 1013
Db      597  TCTGGAATCTCAAGTCCCAAGCTCATGTCCAAAAGCTGATGCTGCTGAGCGCATGTTGT 538
QY      1014  TGGAGATGACAGTGTGCTGTCAGCGCTGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCT 1073
Db      537  GGCTGCTGTTGCTGCTGCTGCTGCGGGGCTCTGAGGTGGAACTGATGCTGCGAGCGCC 478
QY      1074  AGGTAAGCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGC 1433
Db      477  TGTGACGCTCTGACTCAAGTGTGGCGCTGTGACTTTTGCACTTTGCACTTGTGCTGCTGCTGC 418

```

```

QY      1134  TGTGTGAGCGGAGGAGTTCTCCCGCTGCAGCAAGCTTCTATAGGTCATCTCTAAACA 1193
Db      417  TGGTGCATCGCTGGAGCTGCTGCTGTAATCTGCTGTTGCTGCTGCTGGAGCAGG 358
QY      1194  GCCTTACAGACCTCAAAATCTACAGGTTCTGATATATAAAACACAGATATCAGCAGGTTA 1253
Db      357  CTGTGCTGCTGCTGGAGCTGGGCGAGCTGCTGCTGCTGGAGCTGGTCTGCTGCAGCTGA 298
QY      1254  TGATGCTTACAATCCATCAATGATGCTATGTTAGGCGACGAAATGATGCGACTCGTGA 1313
Db      297  TGGAACTGCTGCTGCAGGCGCATGCTGCTGCTGAAAGGCTGGAGCTGCTGCTGAGGCGGA 238
QY      1314  TG 1315
Db      237  TG 236

RESULT 12
LOCUS   BQ424404              890 bp      mRNA      linear      EST 23-MAY-2002
DEFINITION   AGENCOURT_7907155 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154729
5', mRNA sequence.
BQ424404
VERSION      BQ424404.1 GI:21119719
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 890)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13496 row: e column: 02
            High quality sequence stop: 673.
FEATURES    Location/Qualifiers
             source
              1..890
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6154729"
               /tissue_type="retinoblastoma"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH_MGC_67"
               /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
               Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
               Average insert size 1.75 kb. Library constructed by Life
               Technologies."
ORIGIN
Query Match      2.4%; Score 45.2; DB 13; Length 890;
Best Local Similarity 45.3%; Pred. No. 0.67;
Matches 164; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY      954  TCGTGAATCTCAAGTCCCAACAGAGCTCGACACACAGTACAGCAGCAAGCGC 1013
Db      681  TCTGGAATCTCAAGTCCCAAGCTCATGTCCAAAAGCTGATGCTGCTGAGCGCATGTTGT 622
QY      1014  TGGAGATGACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCT 1073
Db      621  GGCTGCTGTTGCTGCTGCTGCTGCGGGGCTGCTGAGGTGGAACTGATGCTGCGAGCGCC 562
QY      1074  AGGTAAGCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGC 1133

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:15 ; Search time 7197.51 Seconds
(without alignments)
11152.641 Million cell updates/sec

Title: US-10-608-559-3
Perfect score: 1852
Sequence: 1 atgagttggcagataagct.....gatcttcaataataaaaga 1852

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_ox.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_in.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1852	100.0	1852	6	AX045133	AX045133 Sequence
2	1848.8	99.8	1256	6	AX045131	AX045131 Sequence
3	1848.8	99.8	10421	1	AE001654	AE001654 Chlamydia
4	1848.8	99.8	12548	1	AE002165	AE002165 Chlamydia
5	1848.8	99.8	300066	1	AE017159	AE017159 Chlamydia
6	1848.8	99.8	300550	1	AP002547	AP002547 Chlamydia
7	1836.2	99.1	110000	6	AX318754_08	Continuation (9 of
8	1748.8	94.4	1956	6	AX338293	AX338293 Sequence
9	1748.8	94.4	1956	6	AX349491	AX349491 Sequence
10	1748.8	94.4	1956	6	AX361992	AX361992 Sequence
11	1165	62.9	2148	1	CHT76KDA	L23921 Chlamydia p
12	1161.8	62.7	2238	6	AX045137	AX045137 Sequence
13	1151	62.1	1456	6	AX045135	AX045135 Sequence
14	361.2	19.5	304769	1	AE016994	AE016994 Chlamydia
15	187.2	10.1	1537	6	BD264444	BD264444 Compounds
16	187.2	10.1	1537	6	AR229257	AR229257 Sequence
17	187.2	10.1	1537	6	AR308168	AR308168 Sequence
18	187.2	10.1	1537	6	AR321875	AR321875 Sequence
19	187.2	10.1	1537	6	AX155814	AX155814 Sequence
20	187.2	10.1	1537	6	AX361664	AX361664 Sequence
21	187	10.1	10954	1	AE001333	AE001333 Chlamydia
22	186	10.0	1944	6	AX786284	AX786284 Sequence
23	185.4	10.0	1941	6	AX362019	AX362019 Sequence
24	184	9.9	1171	6	AX338243	AX338243 Sequence
25	184	9.9	1834	6	AX338268	AX338268 Sequence
26	183	9.9	1983	6	AX338258	AX338258 Sequence
27	148.4	8.0	150	6	AR144067	AR144067 Sequence
28	118	6.4	14698	1	AE002357	AE002357 Chlamydia
29	54.4	2.9	7218	6	I66494	I66494 Sequence (4 of
30	52.4	2.8	77835	2	PFMAL13P2_3	Continuation (24 of
31	52.4	2.8	110000	2	PFMAL13_23	Continuation (24 of
32	51.6	2.8	2205	6	AX317736	AX317736 Sequence
33	50.6	2.7	423	1	MC049	Z33043 M. capricolu
34	49.8	2.7	2000	6	AX655393	AX655393 Sequence
35	49.4	2.7	5098	1	MCU51235	U51235 Mycoplasma
36	49	2.6	1990	6	AR268644	AR268644 Sequence
37	49	2.6	2085	6	AR008322	AR008322 Sequence
38	49	2.6	2085	6	AR027306	AR027306 Sequence
39	49	2.6	2085	6	AR035771	AR035771 Sequence
40	49	2.6	2085	6	AR037631	AR037631 Sequence
41	49	2.6	2085	6	AR078816	AR078816 Sequence
42	49	2.6	2085	6	AR079063	AR079063 Sequence
43	49	2.6	2085	6	AR084296	AR084296 Sequence
44	49	2.6	2085	6	I16740	I16740 Sequence 1
45	49	2.6	2085	6	I70472	I70472 Sequence 1

ALIGNMENTS

RESULT 1
AX045133
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AX045133
Sequence 3 from Patent WO0066739.
AX045133.1
GI:11343732
Chlamydia pneumoniae
Chlamydia pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

Murdin, A.D., Oomen, R.P., Wang, J. and Dunn, P.
i (Chlamydia) antigens and corresponding dna fragments and uses
Thereof
Patent: WO 0066739-A 3 09-NOV-2000;

linear
PAT 24-NOV-2000
0-222/1022

Aventis Pasteur Limited (CA)	
Location/Qualifiers	
1. .1852	
/organism="Chlamydomophila pneumoniae"	
/mol_type="unassigned DNA"	
/db_xref="taxon:83558"	
1. .1752	
/notes="unnamed protein product"	
/codon_start=1	
/transl_table=11	
/protein_id="CAC17378.1"	
/db_xref="GI:11343733"	
/db_xref="REMBL:CAC17378"	
/translation="MSLADKIGIASNSSSSSTSRSDVSTTATATPPPTDDYKT	
QAQAYDTITFTSLADIQAALVSLQDAVINIKDITADDETAIAAEWEKADAVKY	
GAQITELAKYVSNQALIDSLGKLTSLFDLQAALLOSVAANNKAAELKEMQDNVVP	
GKTPAIASLVDDTATQIEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAK	
AAIATATQIAEAOKEKPDSPILQEAQOMVIOAEKDKNKIKPADGPSVNPVPTVGS	
KQGSSTGSIKRVSMLLDAENETASIIKMSGFRMIMHMENTENPDSQAQOELAAQARA	
AKAAGDSAAAALADAQKALEBAALGKAGQOQGINALGQIASRAVSGVPPRAASI	
GSSVKLYKTSKSTGSDYKTIQISAGYDAYKSIDAYGRANDATRDVINNVSTPALTR	
SVPRATEARPEKTDQALRVISGNSRTILGDVYSQVSLQSVNQIITQSNPQANNEI	
RKLTSAVTPKPPQPGYPIVOLSNDSTOKFIKLESIFAEGSRTAABEIKALSPETNSLF	
IQQLVNVIGLSYGLQ"	
ORIGIN	
Query Match 100.0%; Score 1852; DB 6; Length 1852;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 1852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
DB	1 ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
QY	61 TCTGACAGCTGGGATCTCAACGACGCGACCCGACCTAGCGCTCTCTCCACCACGTTTGAT 120
DB	61 TCTGACAGCTGGGATCTCAACGACGCGACCCGACCTAGCGCTCTCTCCACCACGTTTGAT 120
QY	121 GATTATAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACTCAACATCACTAGCT 180
DB	121 GATTATAAGACTCAAGCGCAACAGCTTACGATCTATCTTTACTCAACATCACTAGCT 180
QY	181 GACTATACAGGTGCTTTGGTACGCTCCAGGATGCTGTCACTAATATAAGGATACAGCG 240
DB	181 GACTATACAGGTGCTTTGGTACGCTCCAGGATGCTGTCACTAATATAAGGATACAGCG 240
QY	241 GCTACTGATAGGAAACCGCAATGCTGCGAGTGGGAAAATAAGAAATGCCGATGCGAGTT 300
DB	241 GCTACTGATAGGAAACCGCAATGCTGCGAGTGGGAAAATAAGAAATGCCGATGCGAGTT 300
QY	301 AAAAGTTGGCGCGCAATTTACAGAAATTAGCGAAATATGCTTGGATAAACCAAGCGATTCTT 360
DB	301 AAAAGTTGGCGCGCAATTTACAGAAATTAGCGAAATATGCTTGGATAAACCAAGCGATTCTT 360
QY	361 GACTCTTTAGGTAAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 420
DB	361 GACTCTTTAGGTAAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 420
QY	421 GCACAAATATAACAGAGCTAGCTTCTTAAAGATGCRAGATAACCCAGTAGTCCCA 480
DB	421 GCACAAATATAACAGAGCTAGCTTCTTAAAGATGCRAGATAACCCAGTAGTCCCA 480
QY	481 GGGAAAACCCCTGCAATCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
DB	481 GGGAAAACCCCTGCAATCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
QY	541 ATAGAGAAAGATGGAAATCGGATAGGGATGATATTTTCAGGACAGAACCGCTAGTGA 600
DB	541 ATAGAGAAAGATGGAAATCGGATAGGGATGATATTTTCAGGACAGAACCGCTAGTGA 600
QY	601 GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAACGACGA 660
DB	601 GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAACGACGA 660
QY	661 ATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAATT 720
DB	661 ATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAATT 720
QY	721 CTTCAAGAAAGCGGAACAAATGGTAAATACAGGCTGAGAAAAGATCTTAAAAATATCAAACT 780
DB	721 CTTCAAGAAAGCGGAACAAATGGTAAATACAGGCTGAGAAAAGATCTTAAAAATATCAAACT 780
QY	781 GCAGATGGTTCGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGGAAACAGGA 840
DB	781 GCAGATGGTTCGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGGAAACAGGA 840
QY	841 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTCTGTAGATGATCTCAAAAATAGAGACGCT 900
DB	841 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTCTGTAGATGATCTCAAAAATAGAGACGCT 900
QY	901 TCCATTTTGATGTCGGGTTTCCTCAGATGATTCATGTTTCAATCGGAAAATCCCTGAT 960
DB	901 TCCATTTTGATGTCGGGTTTCCTCAGATGATTCATGTTTCAATCGGAAAATCCCTGAT 960
QY	961 TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAGCTAGACGCGGAAAGCCGCTGGAGAT 1020
DB	961 TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAGCTAGACGCGGAAAGCCGCTGGAGAT 1020
QY	1021 GACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGGCTCTAGGTAAA 1080
DB	1021 GACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGGCTCTAGGTAAA 1080
QY	1081 GCTGGGGCAACAAAGGAGCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
DB	1081 GCTGGGGCAACAAAGGAGCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
QY	1141 AGCGCAGGAGTCTCCCGCTGCGAGCAAGTTCTATAGGTCATCTGTAACACAGCTTTAC 1200
DB	1141 AGCGCAGGAGTCTCCCGCTGCGAGCAAGTTCTATAGGTCATCTGTAACACAGCTTTAC 1200
QY	1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
DB	1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
QY	1261 TACAAATCCATCAATGATGCTATGGTAGGCGCAAGAAATGATGCGACTCGTGTGATGATA 1320
DB	1261 TACAAATCCATCAATGATGCTATGGTAGGCGCAAGAAATGATGCGACTCGTGTGATGATA 1320
QY	1321 AACAAATGTAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA 1380
DB	1321 AACAAATGTAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA 1380
QY	1381 GGACCCAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCTTT 1440
DB	1381 GGACCCAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCTTT 1440
QY	1441 GGAGATGCTTATAGTCAAGTTTCGGCACTCAATCTGTAATGAGATCACTCAGTCCGAAT 1500
DB	1441 GGAGATGCTTATAGTCAAGTTTCGGCACTCAATCTGTAATGAGATCACTCAGTCCGAAT 1500
QY	1501 CCTCAAGCGAATAATAGGAGATCAGACAAAAGCTTACATCGGAGTGACAAAAGCTCCCA 1560
DB	1501 CCTCAAGCGAATAATAGGAGATCAGACAAAAGCTTACATCGGAGTGACAAAAGCTCCCA 1560
QY	1561 CAGTTTGGCTATCCTTATAGTGCAACTTTTCTTAATGACTCTACACAGAAAGTTTCATAGCTAAA 1620
DB	1561 CAGTTTGGCTATCCTTATAGTGCAACTTTTCTTAATGACTCTACACAGAAAGTTTCATAGCTAAA 1620
QY	1621 TTAGAAAGTTTGTGTTGCTGAGGATCTAGGACACGAGCTGAAATATAAGCACTTTCTTT 1680
DB	1621 TTAGAAAGTTTGTGTTGCTGAGGATCTAGGACACGAGCTGAAATATAAGCACTTTCTTT 1680
QY	1681 GAAACGAACTCCTTTGTTTATTTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATTTCTGGT 1740
DB	1681 GAAACGAACTCCTTTGTTTATTTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATTTCTGGT 1740

Qy 1741 TATCTCAATTAACACACCTAAGTGTTCGTTTGGAGAGATTATATATGTGCTTTGGTAAGG 1800
Db 1741 TATCTCAATTAACACACCTAAGTGTTCGTTTGGAGAGATTATATATGTGCTTTGGTAAGG 1800
Qy 1801 CCTTTGTTGAGCCCTTACCAACACACTAGACGATCTTCAATAAATAAAGA 1852
Db 1801 CCTTTGTTGAGCCCTTACCAACACACTAGACGATCTTCAATAAATAAAGA 1852

RESULT 2
AX045131 2156 bp DNA linear PAT 24-NOV-2000
LOCUS
DEFINITION Sequence 1 from Patent WO0066739.
ACCESSION AX045131
VERSION AX045131.1 GI:11343730
KEYWORDS
SOURCE
ORGANISM Chlamydomophila pneumoniae
Chlamydomophila pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.

REFERENCE
1 Murdin, A.D., Oomen, R.P., Wang, J. and Dunn, P.
i(chlamydia) antigens and corresponding dna fragments and uses
thereof
Patent: WO 0066739-A 1 09-NOV-2000;
Aventis Pasteur Limited (CA)
LOCATION/Qualifiers
1. .2156
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
101..2056
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC17329.1"
/db_xref="GI:11343731"
/db_xref="EMBL:CAC17329"
/translation="MVNIGPGPIDETERTPPADLSAQLSASANKSBAEQRINGAE
AKPKESKTSVERNSILSADKSLSSNSSSSSSRSADVSTTATAPT
PPPTDDYKQATYDITFTSLADIQAALVSDQAVTNIKDTAADEETIAAEW
ETKNADAVKGAQITELAKSDNQAILDLSLGLTSLFLLQALILQSVANNKAEELL
KEMONDVVPVGTIPALQSLVDOTDATQIEKQGNARDAVAFQGNASGAVENAKSN
NSISINDAKAIAITAKTQIAEAKKFPDPSPILOEAEQMVIOAEKDLNIKPDGSDV
PNFTGVGSKQGSIGSIRVSMLLDAENETASILMSGFRQMHMENTENPDSQAA
QOELAAQARAAGADSDGAAALADAQALEALGAKGQQQGLNALGQIAASAVVSA
GVPPAAASIGSSVQLYKTSKSDYKQISAGYDAYKSDINDAYGARNDATDVI
NVNVPALTRSPVPRATEARGEKTDQALRVISGNSRTLGDVYQVSLQSVMLIQ
SNPOANNEIRKQLTSVTKPPQFGYPPVQLSNDSTQKFIKLESILFAEGRSRTAEIK
ALSFTNSLFIQQLVNLGSLYGYLQ"

ORIGIN
Query Match 99.8%; Score 1848.8; DB 6; Length 2156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGTCTGGCAGATAAGTGGGTATGCTTCTAGTAAACAGCTCGCTCTCTACTAGCAGA 60
Db 305 ATGAGTCTGGCAGATAAGTGGGTATGCTTCTAGTAAACAGCTCGCTCTCTACTAGCAGA 364
Qy 61 TCTGACAGCTGGACTCAACGACAGCGACCGACCTTACGCTCTCCACCCACGTTTGAT 120
Db 365 TCTGACAGCTGGACTCAACGACAGCGACCGACCTTACGCTCTCCACCCACGTTTGAT 424
Qy 121 GATTATAGACTCAACGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT 180
Db 425 GATTATAGACTCAACGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT 484
Qy 181 GACATACAGCTCTTTGTTGAGCTCCAGGATGCTGTCACTAATAAAGATACAGCG 240
Db 485 GACATACAGCTCTTTGTTGAGCTCCAGGATGCTGTCACTAATAAAGATACAGCG 544
Qy 241 GCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGAAACTAAGAAATGCCATGCGATT 300

Db 545 GCTACTGATGAGGAAACCGCAATCGCTGCGGAGTGGAAACTAAGAAATGCCATGCGATT 604
Qy 301 AAAGTTGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATATAACCAAGCATTCTT 360
Db 605 AAAGTTGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATATAACCAAGCATTCTT 664
Qy 361 GACTCTTTAGTAAACTGACTTCTCTCGACTCTTACAGGCTGCTCTTCTCCAATCTCTGA 420
Db 665 GACTCTTTAGTAAACTGACTTCTCTCGACTCTTACAGGCTGCTCTTCTCCAATCTCTGA 724
Qy 421 GCAAACTAATAACCAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Db 725 GCAAACTAATAACCAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 784
Qy 481 GGGAAACGCGCTGCAATTCCTTCAATCTTTAGTGAACAGATGCTACAGGACACAG 540
Db 785 GGGAAACGCGCTGCAATTCCTTCAATCTTTAGTGAACAGATGCTACAGGACACAG 844
Qy 541 ATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTGCAGGACAGAAACGCTAGTGGGA 600
Db 845 ATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTGCAGGACAGAAACGCTAGTGGGA 904
Qy 601 GCTGTAGAAAAATGCTAAATCTTAATAACAGTATAAGCAACATAGATTTCAGCTAAAGCAGCA 660
Db 905 GCTGTAGAAAAATGCTAAATCTTAATAACAGTATAAGCAACATAGATTTCAGCTAAAGCAGCA 964
Qy 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTTCCCGACTCTCCAATT 720
Db 965 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTTCCCGACTCTCCAATT 1024
Qy 721 CTTCAAGAAAGCGGAACAAATGCTAATACAGGCTCAGAAAGATCTTAAAAATATCAAACT 780
Db 1025 CTTCAAGAAAGCGGAACAAATGCTAATACAGGCTCAGAAAGATCTTAAAAATATCAAACT 1084
Qy 781 GCAGATGCTTCTGATGTTTCCAAATCCAGSAACTACAGTTGGAGGCTCCAAAGCAACAGGA 840
Db 1085 GCAGATGCTTCTGATGTTTCCAAATCCAGSAACTACAGTTGGAGGCTCCAAAGCAACAGGA 1144
Qy 841 AGTAGTATTTGATGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTT 900
Db 1145 AGTAGTATTTGATGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTT 1204
Qy 901 TCCATTTTGTGTTGCTGGGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTT 960
Db 1205 TCCATTTTGTGTTGCTGGGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTTTCCGTT 1264
Qy 961 TCTCAAGCTGCCCAACAGGAGCTCGACACAAGCTAGACAGCGAAAGCGCGCTGGAGAT 1020
Db 1265 TCTCAAGCTGCCCAACAGGAGCTCGACACAAGCTAGACAGCGAAAGCGCGCTGGAGAT 1324
Qy 1021 GACAGTGTGCTGGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1080
Db 1325 GACAGTGTGCTGGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1384
Qy 1081 GCTGGGCAACCAACAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTTTGTG 1140
Db 1385 GCTGGGCAACCAACAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTTTGTG 1444
Qy 1141 AGCGCAGGAGTTCTCCCGCTGACAGCAAGTTCTATAGGGTCACTGTGTAACAGCTTTAC 1200
Db 1445 AGCGCAGGAGTTCTCCCGCTGACAGCAAGTTCTATAGGGTCACTGTGTAACAGCTTTAC 1504
Qy 1201 AAGACCTCAAAATCTTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
Db 1505 AAGACCTCAAAATCTTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT 1564
Qy 1261 TACAATCCATCAATGATGCTTATGGTAGGCGCAAGAAATGATGGACCTCGTGATGTGATA 1320
Db 1565 TACAATCCATCAATGATGCTTATGGTAGGCGCAAGAAATGATGGACCTCGTGATGTGATA 1624
Qy 1321 AACAAATGAAGTACCCCGCTCTCAACAGATCCGTTCCGTAGACACGAAACAGAGCTCGA 1380
Db 1625 AACAAATGAAGTACCCCGCTCTCAACAGATCCGTTCCGTAGACACGAAACAGAGCTCGA 1684

```
1381 GGACCAAGAAAAAAGATCAAGCCCTCGTAGGCTGATTTCTGGCAATAGCAGAACTCTT 1440
Db GGACCAAGAAAAAAGATCAAGCCCTCGTAGGCTGATTTCTGGCAATAGCAGAACTCTT 1744
1441 GGAGATGCTATAGTCAAGTTTCGGCACTCAATCTCTTAATGCAGATCACTCAGTCGAAT 1500
Db GGAGATGCTATAGTCAAGTTTCGGCACTCAATCTCTTAATGCAGATCACTCAGTCGAAT 1804
1501 CCTCAAGCGAATATAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCCCTCCA 1560
Db CCTCAAGCGAATATAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCCCTCCA 1864
1561 CAGTTTGGCTATCTTATGTCACACTTCTTAATGACTCTACACAGAGTTTCATAGCTAAA 1620
Db CAGTTTGGCTATCTTATGTCACACTTCTTAATGACTCTACACAGAGTTTCATAGCTAAA 1924
1621 TTAGAAAGTTTGTTCCTGAAGGATCTAGGACAGCAGCTGAATAAAGACACTTTCCTTT 1680
Db TTAGAAAGTTTGTTCCTGAAGGATCTAGGACAGCAGCTGAATAAAGACACTTTCCTTT 1984
1681 GAAACGAACTCCTTGTATTATTCAGCAGGCTGCTGCTCAATATCGGCTCTCTATATTCGGT 1740
Db GAAACGAACTCCTTGTATTATTCAGCAGGCTGCTGCTCAATATCGGCTCTCTATATTCGGT 2044
1741 TATCTCCAAATCAACACCTAAGTCTTCTGTTGAGAGATTAATTATGCTTTCGTAAGG 1800
Db TATCTCCAAATCAACACCTAAGTCTTCTGTTGAGAGATTAATTATGCTTTCGTAAGG 2104
1801 CCTTGTGTGAGCCCTTACCAACACTAGACGATCTTCAATAATAAAGA 1852
Db CCTTGTGTGAGCCCTTACCAACACTAGACGATCTTCAATAATAAAGA 2156
```

```
RESULT 3
AE001654/1
LOCUS AE001654 Chlamydia pneumoniae section 70 of 103 of the complete genome.
DEFINITION Chlamydia pneumoniae section 70 of 103 of the complete genome.
ACCESSION AE001654
VERSION AE001654.1 GI:4377031
KEYWORDS Chlamydia pneumoniae CWL029
SOURCE Chlamydia pneumoniae CWL029
ORGANISM Chlamydia pneumoniae CWL029
REFERENCE 1 (bases 1 to 10421)
AUTHORS Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R. W.,
Olinger, L., Grimwood, J., Davis, R. W. and Stephens, R. S.
TITLE Comparative genomes of Chlamydia pneumoniae and C. trachomatis
JOURNAL Nat. Genet. 21 (4), 385-389 (1999)
MEDLINE 99206606
PUBMED 10192388
REFERENCE 2 (bases 1 to 10421)
AUTHORS Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Olinger, L.,
Grimwood, J., Davis, R. W. and Stephens, R. S.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1998) Program in Infectious Diseases, University
of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA
FEATURES
Location/Qualifiers
1..10421
/organism="Chlamydia pneumoniae CWL029"
/mol_type="genomic DNA"
/strain="CWL029"
/db_xref="taxon:115713"
133..2751
/gene="CPn0727"
133..2751
/gene="CPn0727"
/codon_start=1
/transl_table=11
/product="C7619 hypothetical protein"
/protein_id="AAD18866.1"
/db_xref="GI:4377032"
/translation="NSTFSGIQNRLRTISGSTRITKLDHKYSGFDRSPVAINLEELN"
```

```
SGIYALRLHMLQSENTNVAAALNPNTIPEPTTSWTDYKHSRPOASSRAPSSORPT
DIVSAAALVLDVGLLVAEVEIDLGALSTISTVROLMA5VLGTLTAEOK
VFFSSYVSEPNLLEHVKQRAAEIOAKQBEIKAVLEAKGVSTREIAILKEYPILY
AADFFKEFTEELHTRYAKVAPIQEMNENAIQLPPTPAITPDNVNENGMNLTSTI
LOADHVLQALQALQGDQIITLQTLVPLVDKTTFKAEFLDITQDLPNTASLKL
YLTDRQIAEYRGKITKVYQNSIQNLSETRVNNRSMLETOISFQAOQNCVFTWIS
QANALNATNKYISAVLTSMEYGLLCLTSYERLADDEKAIEDKSNVEYLPIHI
VVG5WNVNWIAMAAYQELAEYSLGTAVTSODQIKAYLQTRGNEFKATRHPPHNI
QMYFANEITVFNCLTANGAQPDUGGIFIREMTNVGVEADYVNAQRLINEFNTA
ATAHVLQQLQAEIQKQADDDPGKASFTENRKFVAAWITSESLGDLISMLNSQ
LPQEAFLKPLIEEINFNLANALNSLIQTNEFTSTVYSLSYLVSQKGTQNL
AGDYVTLAAAREYIYRDTAKQALNVLNGLKINSPLGATSAQCKENLNTT
YQYSLVTLNQLTVLESLLAGLKMLOTTSNNKYDKSVFKIESFDMPTLALESF
LISGFENISATGGLGPLFTQVQSDQYTSQGTQQLNLNQMTTIQEWTLVSTWQ
VLNGLISQLAGAIYSN"
complement (3078..5033)
/gene="CPn0728"
complement (3078..5033)
/gene="CPn0728"
/codon_start=1
/transl_table=11
/product="CHLPN 76 kDa Homolog_1 (C1622)"
/protein_id="AAD18867.1"
/db_xref="GI:4377033"
/translation="MVNPIGPIDETERTPPADLSAQGLEASAANKSAQRIAGAE
APKESKTSQSVBWSILRSVNALMSLADKLGIASSNSSSTSRSDVDVDTATATP
PPTFDYKTOAGTAYDTITFTSLADIAALVLODAVTNIKDTAETDETAIAEW
ETKNAADVKGAGQITELAKTASNDQALDLSGLKLTSPDLLQALLQSVANNKAAELL
KENQNPVPEKTPAIAQSLVDATQIENDQNAIRDAYFAGQNASGAVENAKSN
NSISNDSAKAAIATATQIAEQKFPDPILOEAPQWIOAEKDKLNKPNDSQV
PNPQTVGGSKQSSGSTRVSMILLDDAENETASILMSGFRQIMHNTENPDQAA
QOELAAQARAAGDDSAALADAKALEAALGKAGQOQILNALQIATAAVVA
GVPAASSTGSSVKOLYKTSKSTQISAGYDAYKSIDAVGARNDATRDVI
NVSTPALTRSPRARTARGPKTDQALARVIGNSRTIGDYVSQALQSVMLIQ
SNPQANEEIRKLTSAVTPKQPGYPIVQLSNDSTQKFIKLESLEPAESRTAAEIK
ALSFTNSLFIQVVLNIGSYGYQ"
complement (5145..6413)
/gene="CPn0729"
complement (5145..6413)
/gene="CPn0729"
/codon_start=1
/transl_table=11
/product="CHLPN 76 kDa Homolog_2 (C1623)"
/protein_id="AAD18868.1"
/db_xref="GI:4377034"
/translation="MKKQVQWLASVVLALTISGYAELPLSEQKVKSHVTYTLDEVK
DYLSKGFVETRKQDGVLRAGDVRARWLYFREDIKNPSDKDKTNPLVNRYSSEYL
YIDYRAERNLSSKMWTAIAGGENTAAGVDINRAFLGYFYKPNPRTDFRMEIGRS
GLGLPFESEVQFSNFDGLHIYMTRELSKDYQVIVHGGPFPVNNMTKKHYAWVEGI
LNRLPKQFVVKCSVVDWNTFVPSSTTEKAAITNAMKYKCVQWMLVGRKHSQVVEIG
QKPLATYGAFLPMPLAKATKTLNGKENLAWFTGTLGLRKGADMSATVRYEVEA
LSVPEIDVSGIGRNLLKFWFAQAIANVDPKEANGFTNYKGFSAWLYMGITDSLSFR
AYGAYSPANDKXGSDFTFRKFDGIISAF"
complement (6516..8159)
/gene="mvn"
/translation="MSRKDNVSLARSIWNILSGTFCSRITGIFREIATAMTYFGADPI
VAAPWLGRTVFFLRKILGLLLEQAFIPHEFRLAQLSRAAFPRFRSERLIKGSLI
IFTLLEAVLWVLYQVEGTYDMILLTPCGIFLMNNVNGALLHCNKFFGVG
LAPVNNIWIFFVIAARHSDPRERIIGLSVALVIGFFHEIITVPGWKLLEAKSP
POEHSVRALLAPLSGLITSSIFQLNLLSDICLARYVHEITGLYMLSKYIQLYLP
LFGFGVTVLLPAISRCVQREDHERGLKMKFLVLTMTSMVIMTAGLILALPGVRV
LYEHLGFPQSAVYAIYVRYGASIIIPMALAPLVSLFYAQROYAVPLTGIGTALA
NIVLSVLGRWLKDVSGISATISATVQVLYFLWYSSKRLPMYKLLWESIRRSIK
VMGTLMACMITLGLNLTQTYVIFLNPPLPPLAWPLISSITAAQAFIASESCIFLAFI
FGFAKLRLVEDLINLASFEYWRGQRLQROHVMQDTQN"
```

```

gene      8143..8319
/ gene="CPn0731"
CDS       8143..8319
/ gene="CPn0731"
/ codon_start=1
/ trans1_table=11
/ product="hypothetical protein"
/ protein_id="AADI8870.1"
/ db_xref="GI:4377036"
/ translation="MSFLLIKFPPPLKQDYLFDISPYTSSEITIGGSYFKLNKASL
QSTLRLRSISIIIS"
gene      8654..9535
/ gene="nfo"
/ note="synonym: CPn0732"
CDS       8654..9535
/ gene="nfo"
/ codon_start=1
/ trans1_table=11
/ product="Endonuclease IV"
/ protein_id="AADI8871.1"
/ db_xref="GI:4377037"
/ translation="MKVLPPPIPLLGHAHTSTAGGLKNAIYEGRDIGASTVQIFIANQ
RWQRALKEVIEDFKALKKEDTDSYIMSHAGYLINCPADPDVILEKSRIGIYEIL
DITLIGISVNPHPGAAKSKEDCNKIVSFOSAPLFDSPPLVLETTAGOGT
LIGSNPEELGYLVONLKNQIPIGVCDVTCIHFAAGYDITSPQGWEDVLEFEYVGLS
YIRAFHLNDSMPELGNKDRHAPLSEGYIGKSFKEFLMDTERKIPKYLETPGPGEN
WQKEIGELKFKSNRDS"
gene      complement(9654..10283)
/ gene="rs4"
/ note="synonym: CPn0733"
CDS       complement(9654..10283)
/ gene="rs4"
/ codon_start=1
/ trans1_table=11
/ product="S4 Ribosomal Protein"
/ protein_id="AADI8872.1"
/ db_xref="GI:4377038"
/ translation="MARYCGPKNVRARREGANIFGRSNPLKKPHPPGHGMORKKK
SDYGLQLEEKQKAKYKAGIMMERQKLVKAFKEVHKGNVAQMFLEFECRLDNMYRM
GFAKTIFAAQQLVAHGHILVNGRRVDRSFFLLRPGMQISLXKSKLQSVKDALESD
ESSLPYSIISLDKTFKFGELLVSPQDQIEAQLPLPINISVVCFFLSHRT"
ORIGIN
Query Match      99.8%; Score 1848.8; DB 1; Length 10421;
Best local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1  ATGAGTCTGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTACTAGCAGA 60
DB      4829 ATGAGTCTGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTACTAGCAGA 4770
QY      61  TCTGCGACGTGGACTCAACGACGACGACGACGACGACGACGACGACGACGACGACGACG 120
DB      4769 TCTGCGACGTGGACTCAACGACGACGACGACGACGACGACGACGACGACGACGACG 4710
QY      121 GATTATAAGACTCAAGCGCAACAGCTTACGATACATCTTTTACCTCAACATCACTAGCT 180
DB      4709 GATTATAAGACTCAAGCGCAACAGCTTACGATACATCTTTTACCTCAACATCACTAGCT 4650
QY      181 GACATACAGGCTGTTTGGTGGAGCTCCAGGATGCTGCTCACTAATATAAAGGATACAGCG 240
DB      4649 GACATACAGGCTGTTTGGTGGAGCTCCAGGATGCTGCTCACTAATATAAAGGATACAGCG 4590
QY      241 GCTACTGATAGGAAACCGGATCGCTCGGAGTGGGAAACATAGAAATGCCGATGCCAGTT 300
DB      4589 GCTACTGATAGGAAACCGGATCGCTCGGAGTGGGAAACATAGAAATGCCGATGCCAGTT 4530
QY      301 AAGCTTCGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATCTTT 360
DB      4529 AAGCTTCGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATCTTT 4470
QY      361 GACTCTCTTAGGTAACATGACTTCCTTCCACCTTTACAGGCTGCTCTTCCCAATCTGTA 420

```

```

DB      4469 GACTCTCTTAGGTAAACTGACTTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 4410
QY      421 GCAAAACAATAACAAAGAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
DB      4409 GCAAAACAATAACAAAGAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 4350
QY      481 GGGAAAACGCTTGCCTCAATCTTCTTAGTTGATCAGACAGATGCTACACGACACAG 540
DB      4349 GGGAAAACGCTTGCCTCAATCTTCTTAGTTGATCAGACAGATGCTACACGACACAG 4290
QY      541 ATAGAGAAGATGGAAATGCGATTAGGATGCATATTTTGCAGGACAGAAACCTAGTGA 600
DB      4289 ATAGAGAAGATGGAAATGCGATTAGGATGCATATTTTGCAGGACAGAAACCTAGTGA 4230
QY      601 GCTGTAGAAAATGCTAAATCTTAATTAACAGTATAAGCAACATAGATTCAGCTAAACAGCA 660
DB      4229 GCTGTAGAAAATGCTAAATCTTAATTAACAGTATAAGCAACATAGATTCAGCTAAACAGCA 4170
QY      661 ATCGCTTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCCAATT 720
DB      4169 ATCGCTTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCCAATT 4110
QY      721 CTTCAAGAAGCGGAACAAATGCTAATACAGGCTGACAAAGATCTTAAAAATATCAAACT 780
DB      4109 CTTCAAGAAGCGGAACAAATGCTAATACAGGCTGACAAAGATCTTAAAAATATCAAACT 4050
QY      781 GCAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAACAGGA 840
DB      4049 GCAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAACAGGA 3990
QY      841 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGATAAATGAGACCGCT 900
DB      3989 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGATAAATGAGACCGCT 3930
QY      901 TCCATTTTGAATGCTGGGGTTTCGTGATGATGATTCATGTTCAATACAGGAAATCTCTGAT 960
DB      3929 TCCATTTTGAATGCTGGGGTTTCGTGATGATGATTCATGTTCAATACAGGAAATCTCTGAT 3870
QY      961 TCTCAGCTGCCCAACAGGAGCTCGCAGCACAGCTAGACGACGCAAGACCGCTGAGAT 1020
DB      3869 TCTCAGCTGCCCAACAGGAGCTCGCAGCACAGCTAGACGACGCAAGACCGCTGAGAT 3810
QY      1021 GCAGTGTCTGTCAGCGCTGCGATGCTCAGAAAGCTTTTGAAGCGGCTCTAGTAAA 1080
DB      3809 GCAGTGTCTGTCAGCGCTGCGATGCTCAGAAAGCTTTTGAAGCGGCTCTAGTAAA 3750
QY      1081 GCTGGGCAACACAGGGGCTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTG 1140
DB      3749 GCTGGGCAACACAGGGGCTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTG 3690
QY      1141 AGCGCAGGAGTTCTCCCGCTGCGACCAAGTTCTTATAGGCTCATCTGTAAAACAGCTTTAC 1200
DB      3689 AGCGCAGGAGTTCTCCCGCTGCGACCAAGTTCTTATAGGCTCATCTGTAAAACAGCTTTAC 3630
QY      1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGAGGTTATGATGCT 1260
DB      3629 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGAGGTTATGATGCT 3570
QY      1261 TACAAATCCATCATGATGCTTATGTTAGGCGACGAAATGATGCGGCTCGTGTGATGATA 1320
DB      3569 TACAAATCCATCATGATGCTTATGTTAGGCGACGAAATGATGCGGCTCGTGTGATGATA 3510
QY      1321 AACAAATGTAAGTACCCCGCTCTCACAGATCGGTTCTTAGAGCACGAAACAGAGCTCGA 1380
DB      3509 AACAAATGTAAGTACCCCGCTCTCACAGATCGGTTCTTAGAGCACGAAACAGAGCTCGA 3450
QY      1381 GGAACGAAAAAACAGATCAAGCCCTCGCTAGGAGTATTTCTGGCAATAGCAGAACTCTT 1440
DB      3449 GGAACGAAAAAACAGATCAAGCCCTCGCTAGGAGTATTTCTGGCAATAGCAGAACTCTT 3390
QY      1441 GGAAGATGCTATAGTCAAGTTTCGGGCACTACAACTGTGTAATGAGATCACTAGTGAAT 1500
DB      3389 GGAAGATGCTATAGTCAAGTTTCGGGCACTACAACTGTGTAATGAGATCACTAGTGAAT 3330

```



```
/gene="CP0016"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37912.1"
/db_xref="GI:7188955"
/translation="MSRKDNVSLARSIFENILSGTFCSRITGIFREIAMAIFYGADPI
VAAPFLGRTVFFKILGGGLLEQAFIPHEFLQAQDLRAAFRRFRSLIKGSTI
IFTLLEAVLVVLYQVEGYDMILLTMIPLPCGIFLMNVNGALLHCENKFFGV
LAPVVNIWIFVFAARHSDPRERIIGLSVALVIGFFEMLIIVPGVVKFLEAKSP
POEHDVALLAPLSLIGLITSSIIFOLNLSLSDICLARYVREIGPLVLMYSLKIQPIH
LEGQVFFVLLPAISRCVOREDEHGLKMKFVLTLTMSVMIIMTAGILLALPGVRV
LYEHGLFQSGAVYALVRVLRGCGASIIIPMALAPLVSVLPFYAQRQYAVPLFYGIGTALA
NIVLSVLGRVWLKVDGSIYATSIAMVQLYFLMYISKRLPMYSLKWLMSIRKSIK
VMGTTMLACMIITGLINILITQTTVYIFLNPFLTFLAWPLSSITTAQAFIUSCIFIPLAFI
FGFAXLLRVEDIILINLASFEYMRGQRLQQRHVMDQTN"
/6108..7376
/gene="CP0017"
/6108..7376
/notes="CP0017"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37913.1"
/db_xref="GI:7188956"
/translation="MKQYQWLASVLLALTYISYALPLSEQVKSHYTTTLDEVK
DYLKRGFVETKQDQVIRIAGDVNRAWLYFREDIKNPSDKDKYNPLRVRYSBFYL
YDYLKRAERWELSKMNTALAGDNTAAGVGNRAFLGVRFKYNPLRDFEWEIGRS
GIGDLFESVQSQSPNDGHIHYWRELSKDPYQVTVHGGPFVNMVTKHYAVWVEGI
LNRFLQFVSVVDWNTVSESTTEKATNANKYKICVQWHLVKGHSQVPLNG
QKKPLYLYGAFNLPLAKATKTLNGENLAWPIGFTGLGLRKGADWSAIVRYEYEA
LSVPEIDVSGIIGRNLKFWFAQAAANYDPPKEANGFTNYKGFSAIYNYGITDSUSFR
AYGAVSKPANDKLGSDFTFRKFDLGIISAF"
/7488..9443
/gene="CP0018"
/7488..9443
/notes="CP0018"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37914.1"
/db_xref="GI:7188957"
/translation="MVPNIPGPGPIDETERTPPADLSAQGLEASAAKSAEARIAGAE
AKPKESKTSVERWSILRGAVNALMSLADKLGIFASSNSSTSRSDVDSTTATATP
PPTFDYKTOAQVDTFTSTSLADIAALVLODAVNIKDTAATDEETAIAAEW
ETKNADAVKGAQITELAXYASDNQALISGLKTSFDLLOALLOSVAANNKAELL
KENDPNVPVPGKTPAQSLVDQTDATQIEKDGNAIRDATFAGQVAGSVAENAKSN
NSISNTSDAKAATATKTOIAEAKKFPDPSILOEAEQWVIOAEKDKNKKPADGSDV
PNPPTVGGSKQGGSDSISIRVSMILLDAENETASTIMSGFRQIMHFNTPDSQAA
QOELAAQAAKAGDGSAAALADAKALAGALGRAGQOGLNLAGQITASAAYVSA
GVPPAASSTGSSVKOLYKTSKSTGDKYTOISAGDYAKSINDYGRANDATRVII
NVVSTALTRSPRATREARPEKTDQALARVISGNSRTLGDYVSQVLSALQVWQIIV
SNPQANNEIRKQTSATVTPQFGYPYVOLSNDSTQKFKAKUESLFAESRTAAABIK
ALSFETNSLFIQQLVNLIGLSYGIQ"
/complement(9770..12388)
/gene="CP0019"
/complement(9770..12388)
/notes="CP0019"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37915.1"
/db_xref="GI:7188958"
/translation="MSTFSIQNRLRTISGESRIIKLDHKYSGFDRSPVPAINLEELN
SGIYALRHLNALQSENTVAALLNPNTIFPTTSWTDYKHSRQASSPRAPSSQTP
```

```
DIVSAAALALVLVIDGGIAELVASVTEIDLGALSTISTVRQLMASYILGTLTTAEQEK
VVFSSVVPSEKLNLEHVQKFAEIOAKOBEIKAVLEAKGVSTEEBAIKLEYDPIY
AADPFKEFIEBPLHTYPAKGAPIQEMNENAIQLLPATTPNPNVNEVGNATYSTI
LOADDDAIKOAPALGGQBEIITILQTLVPLVDKTTFFKAEFDLTYTQLPNTASLKL
YLTDRQIAEYRKITKVYQNSIQNSETSRVVRNNRSMLETQKLSMFQOQNCFTWIS
QANALNATNKYISAVLTTSMEYMGGLLSUSMYERLADDEKKAIFDKSVNEYLPDHI
VVGSGWNGMTAKMAAYQELAYSLGTAVTSDQIKAYLQTRGNEFKATRHHFNIGD
QMYOFANETVFGNCLTTANGAIOPDLGGFIREAMTNGVTVEADYVSNARQIRINEFNTA
ATAHVQLQLQIAELQKKADDLPDKASFTENRKXFAVAAMITSESLGDLISMLINSQ
Query Match 99.8%; Score 1848.8; DB 1; Length 12548;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGTCTGGCAGATAAGCTGGGTATTCCTCTAGTAACAGCTCGCTCTTCTACTAGCAGA 60
DB 7692 ATGAGTCTGGCAGATAAGCTGGGTATTCCTCTAGTAACAGCTCGCTCTTCTACTAGCAGA 7751
QY 61 TCTGCAGACGTGGACTCAACGACGCGACCTACGCTCTCTCCACCCACGCTTTGAT 120
DB 7752 TCTGCAGACGTGGACTCAACGACGCGACCTACGCTCTCTCCACCCACGCTTTGAT 7811
QY 121 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTTACCTCAACATCAGTACT 180
DB 7812 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTTACCTCAACATCAGTACT 7871
QY 181 GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGCATTAATATAAGGATACAGCG 240
DB 7872 GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGCATTAATATAAGGATACAGCG 7931
QY 241 GCTACTGATGAGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGAGTT 300
DB 7932 GCTACTGATGAGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGAGTT 7991
QY 301 AAAGTTGGCGGCCAAATTACAGAAATAGCGAAATATGCTTCGGGATAACCAAGCATTCCT 360
DB 7992 AAAGTTGGCGGCCAAATTACAGAAATAGCGAAATATGCTTCGGGATAACCAAGCATTCCT 8051
QY 361 GACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 420
DB 8052 GACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 8111
QY 421 GCAACAATTAACAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA 480
DB 8112 GCAACAATTAACAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA 8171
QY 481 GGGAAACCCCTGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACACAG 540
DB 8172 GGGAAACCCCTGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACACAG 8231
QY 541 ATAGAGAAAGATGGAAATCGGATTAGGGATGCATATTTTGCAGGACAGAACGCTAGTGA 600
DB 8232 ATAGAGAAAGATGGAAATCGGATTAGGGATGCATATTTTGCAGGACAGAACGCTAGTGA 8291
QY 601 GCTGTAGAAAAATGCTAAATCTTAATAACAGTATAAGCAATAGATTACACTAAGCAGCA 660
DB 8292 GCTGTAGAAAAATGCTAAATCTTAATAACAGTATAAGCAATAGATTACACTAAGCAGCA 8351
QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTCCCGCAGCTCTCCAATT 720
DB 8352 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTCCCGCAGCTCTCCAATT 8411
QY 721 CTTCAAGAAAGCGGAACAAATGGTAATACAGGCTGAGAAAAAGATCTTAAAAAATATCAAACCT 780
DB 8412 CTTCAAGAAAGCGGAACAAATGGTAATACAGGCTGAGAAAAAGATCTTAAAAAATATCAAACCT 8471
QY 781 GCAGATGGTTCGATGTTCCAAATCCAGAACTACAGTTCAGTTGGAGGCTCCCAAGCAACAGGA 840
DB 8472 GCAGATGGTTCGATGTTCCAAATCCAGAACTACAGTTCAGTTGGAGGCTCCCAAGCAACAGGA 8531
QY 841 AGTAGTATTGGTAGTATTGTTCCATGCTGTTAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 8532 AGTAGTATTGGTAGTATTGTTCCATGCTGTTAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 9591
```

KEYWORDS		Chlamydia pneumoniae TW-183
SOURCE		Chlamydia pneumoniae TW-183
ORGANISM		Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia
REFERENCE		1 (bases 1 to 300066)
AUTHORS		Geng, M.M., Schubmacher, A., Muehldorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T., Essig, A., Marre, R. and Melchers, K.
TITLE		The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis
JOURNAL		Unpublished
REFERENCE		2 (bases 1 to 300066)
AUTHORS		Geng, M.M., Schubmacher, A., Muehldorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T. and Melchers, K.
TITLE		Direct Submission
JOURNAL		Submitted (28-MAY-2002) RDR/IT, RDR/FG, RDR/BT, ALTANA Pharma, Byk-Gulden-Str. 2, Konstanz 78467, Germany
FEATURES		Location/Qualifiers
source		1..300066
		/organism="Chlamydia pneumoniae TW-183"
		/mol_type="genomic DNA"
		/strain="TW-183"
		/db_xref="taxon:182082"
		/note="type strain"
		biovar: TWAR
		126..797
		/locus_tag="CpB0543"
		126..797
		/locus_tag="CpB0543"
		/notes="similar to hypothetical protein CT433 - Chlamydia trachomatis"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="AAP98472.1"
		/db_xref="GI:33236384"
		translation="MTLYGLNQKTKARKYQAHYLPILTFFPYAKSTQNKRALQFLPQ ATHVILTPSSSTHLFSRMTSLSKATKTKYLCIGESTKERLLSLGQVYVATQ EAGIFPILQALPSSARILYPHSSILARVIRFLYRNFTFFSYHYTVKPKLKNKI LSYKXILLTSPSTVRAFAKIFPRFEKTYWCQGRMTLQBFQFSQKQVSLLETGLK SRVSP"
		862..1194
		/locus_tag="CpB0544"
		862..1194
		/locus_tag="CpB0544"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="AAP98473.1"
		/db_xref="GI:33236385"
		translation="MASSATPGDGTAPSLFPLATPRYNFKLALFVITIALVWIAL IATVATGLCHPLCSFIFLTPFISRYICSHYARNVYIALDVVPDHSKLQDMRS HSFIFSDR"
		1242..2321
		/locus_tag="CpB0545"
		1242..2321
		/locus_tag="CpB0545"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="AAP98474.1"
		/db_xref="GI:33236386"
		translation="MSGPSRSTESSQVSLSYPRDKIAPKQFTIAKISTLAILASL ALGALVAGISLTIVLGNPVFLALLITTFALSVVTFVYHMTSKVSNWOKVLEQNFK PLGAWQEKXNDVCSNEMQYNNHINAPKFKVAICTDASQFPQFTLGLVIEQNST GIIFNPVGNLINDTAINSLIYSLTKDKSVNDCKQREGPAGKEDPSPTEVAV KALENEALDDTFLNLSAASKSILPFLGHVCGPKSEELPNQCEYTRQALLAYENCL KAALESAAVALPFTSVYVPEPEILPEKGTGYWNTQTAFCCKRALLDAIQNTAUR YQPSRLVLQDPENTIESQSRSE"
		2661..3425
		/locus_tag="CpB0546"
		2661..3425
		/locus_tag="CpB0546"

RESULT 5

AE017159/c

LOCUS

DEFINITION

Chlamydia pneumoniae TW-183, section 3 of 4 of the complete genome.

ACCESSION

AE017159

VERSION

AE017159.1

GI:33236383

300066 bp

DNA

linear

BC1 25-JUL-2003


QY	301	AAAGTTGGCGCGAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGGATCTTT	360	1381	GGACCAGAAAAACAGATCAAGCCCTCGTAGGTGATTTCTGGCAATAGCAGACTCTT	1440
Db	214097	AAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGGATCTTT	214038	213017	GGACCAGAAAAACAGATCAAGCCCTCGTAGGTGATTTCTGGCAATAGCAGACTCTT	212958
QY	361	GACTCTTTAGGTAAATGATGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA	420	1441	GGAGATGCTATAGTCAAGTTTTCGGCACTACAAATCTGTAATCGAGATCACTCAGTCGAAT	1500
Db	214037	GACTCTTTAGGTAAATGATGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA	213978	212957	GGAGATGCTATAGTCAAGTTTTCGGCACTACAAATCTGTAATCGAGATCACTCAGTCGAAT	212898
QY	421	GCAAAACAATAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGTAAACCCAGTAGTCCCA	480	1501	CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGCGAGTGACAAAGCTTCCA	1560
Db	213977	GCAAAACAATAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGTAAACCCAGTAGTCCCA	213918	212897	CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGCGAGTGACAAAGCTTCCA	212838
QY	481	GGGAAACGCGCTGCAATGCTCAATCTTATAGTTGATCAGACAGATGCTACAGCGACAG	540	1561	CAGTTTGGCTATCTTATGTCGCACTTCTTAATGACTCTTACACAGAAGTTATAGTAAA	1620
Db	213917	GGGAAACGCGCTGCAATGCTCAATCTTATAGTTGATCAGACAGATGCTACAGCGACAG	213858	212837	CAGTTTGGCTATCTTATGTCGCACTTCTTAATGACTCTTACACAGAAGTTATAGTAAA	212778
QY	541	ATAGAGAAAGATGGAATGCGATTAGGGATGCAATATTTTCGAGGACAGACGCTAGTGA	600	1621	TTAGAAAAGTTTGTTCGCTGAAGGATCTAGGACAGCAGCTGAAATAAAAGCACTTCTCTTT	1680
Db	213857	ATAGAGAAAGATGGAATGCGATTAGGGATGCAATATTTTCGAGGACAGACGCTAGTGA	213798	212777	TTAGAAAAGTTTGTTCGCTGAAGGATCTAGGACAGCAGCTGAAATAAAAGCACTTCTCTTT	212718
QY	601	GCTGTAGAAAATGCTAAATCTTAATAAGCAATAAGCAACATAGATTACAGTAAAGCAGCA	660	1681	GAAACGAATCTCTTGTATTATTCAGCAGGTGCTGTCGAATATCGGCTCTATATCTTGGT	1740
Db	213797	GCTGTAGAAAATGCTAAATCTTAATAAGCAATAAGCAACATAGATTACAGTAAAGCAGCA	213738	212717	GAAACGAATCTCTTGTATTATTCAGCAGGTGCTGTCGAATATCGGCTCTATATCTTGGT	212658
QY	661	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAAAT	720	1741	TATCTCCAATAACAACACCTTAAGTGTTCGTTGGAGAGATTATATGCTTTGGTAAGG	1800
Db	213737	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAAAT	213678	212657	TATCTCCAATAACAACACCTTAAGTGTTCGTTGGAGAGATTATATGCTTTGGTAAGG	212598
QY	721	CTTCAGAGCGGAACAATGTAATACAGCGTGAAGAAAGATCTTAAATAATCAAACT	780	1801	CCTTTGTTGAGGCTTACCACACACACTAGAACGATCTTCAATAATAAAGA	1852
Db	213677	CTTCAGAGCGGAACAATGTAATACAGCGTGAAGAAAGATCTTAAATAATCAAACT	213618	212597	CCTTTGTTGAGGCTTACCACACACACTAGAACGATCTTCAATAATAAAGA	212546
QY	781	GCAGATGTTCTGATGTTCCAAATCCAGGAACACTACAGTTGGAGGCTCCAAAGCAACAGGA	840	RESULT 6		
Db	213617	GCAGATGTTCTGATGTTCCAAATCCAGGAACACTACAGTTGGAGGCTCCAAAGCAACAGGA	213558	AP002547/c		
QY	841	AGTAGTATTGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT	900	LOCUS	300550 bp	DNA
Db	213557	AGTAGTATTGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT	213498	DEFINITION	Chlamydia pneumoniae J138 genomic DNA, complete sequence,	linear BCT 25-APR-2003
QY	901	TCCATTTTGTGCTGCGGTTTCGTCAGATGATTCACATGTTTCAATAGGAAATCTCTGAT	960	ACCESSION	AP002547	section 3/4
Db	213497	TCCATTTTGTGCTGCGGTTTCGTCAGATGATTCACATGTTTCAATAGGAAATCTCTGAT	213438	VERSION	AP002547.2	GI:10176693
QY	961	TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAAGCTAGACAGCGAAGCGCTGGAGAT	1020	KEYWORDS	Chlamydia pneumoniae J138	
Db	213437	TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAAGCTAGACAGCGAAGCGCTGGAGAT	213378	SOURCE	Chlamydia pneumoniae J138	
QY	1021	GACAGTGTGCTGAGCGCTGGAGATGCTCAGAAAAGCTTTAGAACGGCTCTAGGTAAA	1080	ORGANISM	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.	
Db	213377	GACAGTGTGCTGAGCGCTGGAGATGCTCAGAAAAGCTTTAGAACGGCTCTAGGTAAA	213318	REFERENCE	Shirai, M., Hirakawa, H., Ouchi, K., Tabuchi, M., Kishi, F., Kimoto, M.,	
QY	1081	GCTGGCAACAACAGGCTATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTTGTG	1140	AUTHORS	Takeuchi, A., Nishida, J., Shibata, K., Fujinaga, R., Yoneda, H.,	
Db	213317	GCTGGCAACAACAGGCTATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTTGTG	213258	TITLE	Natsushima, H., Tanaka, C., Furukawa, S., Miura, K., Nakazawa, A.,	
QY	1141	AGCGCAGAGTTCTCCGCTGCGAGAGTTCTATAGGCTCATCTGTAACACGCTTAC	1200	JOURNAL	Ishii, K., Shiba, T., Hattori, M., Kuhara, S. and Nakazawa, T.	
Db	213257	AGCGCAGAGTTCTCCGCTGCGAGAGTTCTATAGGCTCATCTGTAACACGCTTAC	213198	MEDLINE	Comparison of outer membrane protein genes omp and pmp in the whole	
QY	1201	AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT	1260	PUBMED	genome sequences of Chlamydia pneumoniae isolates from Japan and	
Db	213197	AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT	213138	REFERENCE	the United States	
QY	1261	TACAAATCCATCAATGATGCTTATGAGGACGCAAAATGATGCGACTCGTGATGTGATA	1320	AUTHORS	J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)	
Db	213137	TACAAATCCATCAATGATGCTTATGAGGACGCAAAATGATGCGACTCGTGATGTGATA	213078	TITLE	Shirai, M.	
QY	1321	AACATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGCTCGA	1380	JOURNAL	Direct Submission	
Db	213077	AACATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGCTCGA	213018	REFERENCE	Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University	

```

FEATURES
    source          Location/Qualifiers
1. .300550
    /organism="Chlamydomydia pneumoniae J138"
    /mol_type="genomic DNA"
    /strain="J138"
    /db_xref="taxon:138677"
    /note="synonym:Chlamydia pneumoniae (strain J138)"
    complement(21. .794)
    /gene="dapF"
    complement(21. .794)
    /gene="dapF"
    /codon_start=1
    /transl_table=11
    /product="diaminopimelate epimerase"
    /protein_id="BAA98725.1"
    /db_xref="GI:8978890"
    /translation="NAFYSPSTISKYFIYSGAGNRFLGLGETLPEVEDVRFLCQETRVDD
    GFLYKPPSCADAQLIIFNSDGSRRPFCMCGNLCRAIAHLASOKGSDISVSDGLYS
    GYFYSDRVLVMTADWASVHRLESRPDLPEKVCVCHTGVPHAVILPEISLIDL
    SILGPELRYHOTFSPDGVNVFVQLLGHQCLRVRYVERGVEGTAACTGALASALV
    SNSYGWKESIHTWGGELMTVSQNRGVLLQSVTRDL"
    complement(763. .1338)
    /gene="clpP.1"
    complement(763. .1338)
    /gene="clpP.1"
    /codon_start=1
    /transl_table=11
    /product="Clp protease"
    /protein_id="BAA98726.1"
    /db_xref="GI:8978891"
    /translation="MADGSRVHKLRIIEKELLEARRVFSEPVTEKSASDAIKLWYL
    ELKDGKPIVFVNSPGSVDAQFAVMDIKMLTSPVTVTGLAASGVSLSLAAP
    GRFPATPSRIMHQPSIGSGPTGQATDLDIHAREILTKARLIIDVYEATNQPRDII
    EXAIDRDMWTANEAQDFLLGILFSFNDL"
    complement(1358. .2917)
    /gene="glyA"
    complement(1358. .2917)
    /gene="glyA"
    /codon_start=1
    /transl_table=11
    /product="serine hydroxymethyltransferase"
    /protein_id="BAA98727.1"
    /db_xref="GI:8978892"
    /translation="MLKVPEKPKKFAIVEIFTKVAVVSVLLHKFLFNASGKKQSILAS
    TAYLAALDHLNAPSGIRIDELKQSRSHLXIASENYSLSVGLANGNLLTKYC
    EGSPFRFYSCCNDAIEWCVEVETAKELPAACQVPHSGADANLLAVMALITHKV
    QGPAVSKLGTVNEETEEYTLKAEKSSVCCLGSLNSGGHLLTHGNVRLNWSKLM
    RCFPVNPDTECFDAEISRLAKYKPKYLAGYSYGRRLNFAYLKQIAEDCGSLV
    WYDMAFAGLVAGGVVDENPIPYADIVTTTHKTLRGRGGLVLAETREYSLNKA
    PLMMGSGPLPHVIAAKVALKEALSDFDKYAHOVVNARRLAERFLSHGLITLGGT
    DNHMVYDGLSGISKIADILLSSVGIARNRSLPSDAIGKWDTSGLRGTALTTL
    GWDIDMEVADIIVKVLNRILRSLCHVEGSSKKNKGLPEATAQEARDRVRLNLLRFP
    LYPEIDLEALV"
    3107. .3778
    /gene="CPj0522"
    3107. .3778
    /gene="CPj0522"
    /codon_start=1
    /transl_table=11
    /product="CT433 hypothetical protein"
    /protein_id="BAA98728.1"
    /db_xref="GI:8978893"
    /translation="MTLYGLNOKTARKQAQHYLPILTLPFYAKSTPQNKRALQFLPQ
    ATHVILTSPSSTHLFSRMTSLLSKATLTKTYLCTGISTEKRLLSFLGQVKVVAIQ
    IEAEGIFFDLQALPVSILRPHSSILARPVIREFLYNRFTFFSYPHYTKPRKKNNI
    LSKYKXIIFTSPTVRAFAKIFRPPEKTYCQGRMTLQEFQFESSQKQVSLLETIGK
    SRTSp"
    3843. .4175
    /gene="CPj0523"
    3843. .4175
    /gene="CPj0523"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="BAA98729.1"
    /db_xref="GI:8978894"
    /translation="MASSATFGDGTAPSLFPFPATPRYNFKLALFVTIATLAWIAL
    IATTIAGLCIHLCLCSFILTALIFYISRYCISHYARNVYIALDVVPDHSKLODMRS
    HSPFIASDR"
    4223. .5302
    /gene="CPj0524"
    4223. .5302
    /gene="CPj0524"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="BAA98730.1"
    /db_xref="GI:8978895"
    /translation="MSGFSRTESQSVLSVPRKETAIPKKQFTIAKISTAILLASL
    ALGALVAGISITVILGNFVFLALILFVVTFLVHQTSKVSSNWQKVEQNFK
    PLGKAWQCNQVSNEMFYNNHLPKFAIQTADSQPQFTPLTGLRVLEKQST
    GIIFNVPQPTNLIDNTATNLSTLYTLKDKSVMDTKQREGPAKGDPPSPTEVRV
    VKLPNEALDQTNLNLSSAEKSIILPTFLGHVCGPKSELPNQQYVYRQALLAYENCL
    KAAIESHAALVALPLFTSVYVPPPEILLPKGSTFYWDNQTAQACREALLDALQNTALR
    IQRSLVLQDPFNLTIESQSRSE"
    5642. .6406
    /gene="CPj0525"
    5642. .6406
    /gene="CPj0525"
    /codon_start=1
    /transl_table=11
    /product="CT398 hypothetical protein"
    /protein_id="BAA98731.1"
    /db_xref="GI:8978896"
    /translation="MHDALLSLTAQLDIDKIMRLMRVKKEHOKELAKVQSLKSDIRR
    KYQKELMENLKTQIRGENRIQELISQINKLENQQAQKMDENALTOEMWTANK
    ERSEHOLSDMRQAGEDLIIVLSKESLASTENSSSVIEKEIFESIKKINEBKAL
    LEQRTLKHATNPELLISYERLNNKKORVVPIENRVCSGCHI VLTPOHENVLRKDO
    RLIFCEHCSRILYQWESQVNAQENSTAKRRRRRAAV"
    complement(6827. .7816)
    /gene="kpsP"
    complement(6827. .7816)
    /gene="kpsP"
    /codon_start=1
    /transl_table=11
    /product="GutQ/KpsF family sugar-P isomerase"
    /protein_id="BAA98732.1"
    /db_xref="GI:8978897"
    /translation="NPSPMISDVCQDILKQKEAVDFFQAFQPKQKAMQLAELKILGH
    SGWFFSGVGKSGCVARKLAVATLQLSERALFFSPVLLHGLDGLVSPGDI VCLFSSKS
    GETQLLDTVPLKSRRAILVAITSMPSYNSLAALSDLVILPSVAELDFNLIPNTST
    TQMIIFGDFLAMLHFSRGSLSLYGKHPSCQVGMKANGKVDPMFPKTEYVFFCHLG
    DKVSTSLVFSAYGCCVCIVDQFRLMGIFTDGLRSLASGYGGEVLSLEKVMTA
    NRCITSDSDIALQLMESSPPAVLPVLVDNEENRHTVTGLLHMTLAKAGLL"
    complement(7843. .9024)
    /gene="sucB.2"
    complement(7843. .9024)
    /gene="sucB.2"
    /codon_start=1
    /transl_table=11
    /product="dihydrolipoamide succinyltransferase"
    /protein_id="BAA98733.1"
    /db_xref="GI:8978898"
    /translation="MIFFRPPKIGTSSGGSIVRWLNKLGHVARDPELIVSTDKI
    ATELPSPKAGRLVRCFVNEGDEVASDGLVLELEI SEADDESTCPLTSCETKSEA
    GSSSSVWFSPAVLSLAQREGIGLDNLQKIAGTGGKRVTRQDLRLAYISESQVSIPE
    IQQEVNRI PMSPLRAIASSLSKSDVEPHASIVVDVDTLMLNLSGERQFLDTH
    GVKYLTFTIVOCQAQLRQFFLLNGSLDGTIIVMKKSVNAGVAVNLAKGQVVPVIVH
    NQDQRLYSIAKALADLSRAKLNKLDPEVDQSVTVTNFGMTGALIGMPILIRYFEV
    ALIGTGTIQRVVRDDPSLAIRKMVYVTLTFDHRVLDDIYGSEFLTSLKKNLESVTM
    G"
    complement(9038. .10282)
    /gene="gltT"
    complement(9038. .10282)
    /gene="gltT"
    /codon_start=1

```

gene
CDS



gene
CDS

gene
CDS

gene
CDS

gene
CDS

```
/transl table=11
/product="glutamate symport"
/protein_id="BA998734.1"
/db_xref="GI:897889"
/translation="MKLWMKIFGLFVGVTGLGLVLEDKAIFPKPIGDIPLNLSMVVY
PLVFCSMVLGASISDMKUGRIGIKSVGYLGTALAI VIGLCFAWIFPFGNGCDPA
QAQMSDAVTVISNKAAYFLSIIAQVFPNVRPSFAENILQIIFAIPLGLAURL
SGRGRVERPFDGFSILMRNMVINSFAPYGVGSAWMI SGNHGLGVLMWLGKFI
AYVLACLFHATLVFGGLVRFCCKMSFKLSMMDAISCAVSTASSATPVTMRCVVS
KNLGVSEVSGFVLPLGATVNMGTALFQGMAAVFIQAQVNCPLSLSLLLVVTAIF
VAGSAGVPGGMIILGSULASVGLPIQGTAILAGIDRLRIDVGTPTNLIIGDAVATY
SAGSELSPPGSEMIKQESVET"
gene complement (10282..11379)
/ gene="ycah"
CDS complement (10282..11379)
/ gene="ycah"

Query Match 99.8%; Score 1848.8; DB 1; Length 300550;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 60
DB 217394 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 217335

QY 61 TCTCAGACGTGGACTCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 120
DB 217334 TCTCAGACGTGGACTCAACGACGACGACGACGACGACGACGACGACGACGACGACGACG 217275

QY 121 GATTATAGACTCAGCGCAACAGCTTACGATCTATCTTTTACCTCAACATCACTAGCT 180
DB 217274 GATTATAGACTCAGCGCAACAGCTTACGATCTATCTTTTACCTCAACATCACTAGCT 217215

QY 181 GACATACAGGCTGCTTTGGTAGGCTCCAGGATGCTGTCACATAATATAAGGATACAGCG 240
DB 217214 GACATACAGGCTGCTTTGGTAGGCTCCAGGATGCTGTCACATAATATAAGGATACAGCG 217155

QY 241 GCTACTGATAGGAAACCGCAATCGCTGGGAGTGGGAAACATAAGATGCCGATGCAGTT 300
DB 217154 GCTACTGATAGGAAACCGCAATCGCTGGGAGTGGGAAACATAAGATGCCGATGCAGTT 217095

QY 301 AAAGTTGGCGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 360
DB 217094 AAAGTTGGCGCGCAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 217035

QY 361 GACTCTTTAGGTAAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 420
DB 217034 GACTCTTTAGGTAAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 216975

QY 421 GCAACAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
DB 216974 GCAACAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 216915

QY 481 GGGAAAAACGCTGCAATGCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGACACAG 540
DB 216914 GGGAAAAACGCTGCAATGCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 216855

QY 541 ATAGAGAAAGATGGAAATCGAATTAGGATGATATTTTGCAGGACAGAAACGCTAGTGA 600
DB 216854 ATAGAGAAAGATGGAAATCGAATTAGGATGATATTTTGCAGGACAGAAACGCTAGTGA 216795

QY 601 GCTGTAGAAATGCTTAAATCTTAATACAGTATAAGCAACATAGATTACGTAAGCAGCA 660
DB 216794 GCTGTAGAAATGCTTAAATCTTAATACAGTATAAGCAACATAGATTACGTAAGCAGCA 216735

QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCGCCGACTCTCCAATT 720
DB 216734 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCGCCGACTCTCCAATT 216675

QY 721 CTTCAAGAAACGCGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTTAAAAAATATCAAACT 780
DB 216674 CTTCAAGAAACGCGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTTAAAAAATATCAAACT 216615
```

```
QY 781 GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACAGGA 840
DB 216614 GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACAGGA 216555

QY 841 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTGTAGATGATGCTGCTGAAATAGACCGCT 900
DB 216554 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTGTAGATGATGCTGCTGAAATAGACCGCT 216495

QY 901 TCCATTTTGTATGCTCTGGTTTCTCAGATGATTCATATGTTTCAATACGGAATAATCTTGAT 960
DB 216494 TCCATTTTGTATGCTCTGGTTTCTCAGATGATTCATATGTTTCAATACGGAATAATCTTGAT 216435

QY 961 TCTCAAGCTGCTCCAAACAGGAGCTCGCAGCACAGCTAGACGCGGAAACGCTGGAGAT 1020
DB 216434 TCTCAAGCTGCTCCAAACAGGAGCTCGCAGCACAGCTAGACGCGGAAACGCTGGAGAT 216375

QY 1021 GACAGTGTCTGCTGAGCGCTGGCAGATGCTTCAGAAAAAGCTTTAGAAAGCGGCTCTAGGTAAA 1080
DB 216374 GACAGTGTCTGCTGAGCGCTGGCAGATGCTTCAGAAAAAGCTTTAGAAAGCGGCTCTAGGTAAA 216315

QY 1081 GCTGGGCAACAAACAGGCACTATCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
DB 216314 GCTGGGCAACAAACAGGCACTATCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 216255

QY 1141 AGCGCAGAGTTCCTCCCGCTGCGACGAAAGTTCTATAGGGTCATCTGTAAAAACAGCTTTAC 1200
DB 216254 AGCGCAGAGTTCCTCCCGCTGCGACGAAAGTTCTATAGGGTCATCTGTAAAAACAGCTTTAC 216195

QY 1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGCT 1260
DB 216194 AAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGCT 216135

QY 1261 TACAAATCCATCAATGATGCTTATGTTAGGACGAAATGATGCGACTCGTGATGCTATA 1320
DB 216134 TACAAATCCATCAATGATGCTTATGTTAGGACGAAATGATGCGACTCGTGATGCTATA 216075

QY 1321 AACAACTGTAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA 1380
DB 216074 AACAACTGTAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA 216015

QY 1381 GGACAGAAAAAACAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATAGCAGAACTCTTT 1440
DB 216014 GGACAGAAAAAACAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATAGCAGAACTCTTT 215955

QY 1441 GGAGATGCTATAGTCAAGTTTGGCACTACAACTCTTAATGAGATGACATCTCACTCCGAT 1500
DB 215954 GGAGATGCTATAGTCAAGTTTGGCACTACAACTCTTAATGAGATGACATCTCACTCCGAT 215895

QY 1501 CCTCAAGCGAATAATGAGGAGATCAGACAAAGCTTACATCGGCAGTGACAAAGCCTCCA 1560
DB 215894 CCTCAAGCGAATAATGAGGAGATCAGACAAAGCTTACATCGGCAGTGACAAAGCCTCCA 215835

QY 1561 CAGTTTGGCTATCTTATGTGCAACTTTCTTAATGACTCTACACAGAGTTTCATAGCTAAA 1620
DB 215834 CAGTTTGGCTATCTTATGTGCAACTTTCTTAATGACTCTACACAGAGTTTCATAGCTAAA 215775

QY 1621 TTAGAAAGTTTGTGTTGCTGAGAGATCTAGGACAGCTGAAATATAAAGCACTTTCCTTT 1680
DB 215774 TTAGAAAGTTTGTGTTGCTGAGAGATCTAGGACAGCTGAAATATAAAGCACTTTCCTTT 215715

QY 1681 GAAACGAACTCCTTGTGTTTATTTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATCTGGT 1740
DB 215714 GAAACGAACTCCTTGTGTTTATTTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATCTGGT 215655

QY 1741 TATCTCAATTAACAAACCTTAAGTGTGTTTGGAGAGATTAATTATGCTGCTTGGTGAAG 1800
DB 215654 TATCTCAATTAACAAACCTTAAGTGTGTTTGGAGAGATTAATTATGCTGCTTGGTGAAG 215595

QY 1801 CCTTGTGTTGGGCTTACCAACACTAGACGATCTTCAATAAATAAAGA 1852
DB 215594 CCTTGTGTTGGGCTTACCAACACTAGACGATCTTCAATAAATAAAGA 215543
```


Sequence split into 13 fragments LOCUS AR310754 Accession AR310754															
RESULT 7	Fragment Name	Begin	End												
AR310754_08/c	AR310754_00	1	110000												
WPCOMMENT	AR310754_01	100001	210000												
Sequence split into 13 fragments LOCUS AR310754 Accession AR310754															
Query Match 99.1%; Score 1836.2; DB 6; Length 110000;															
Best Local Similarity 99.8%; Pred. No. 0;															
Matches 1849; Conservative 0; Mismatches 3; Indels 1; Gaps 1;															
QY	1	ATGAGTCTGGCAGATAAGCTGGGTATTCCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	60												
DB	28394	ATGAGTCTGGCAGATAAGCTGGGTATTCCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	28335												
QY	61	TCTGCAGACGTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTTTGAT	120												
DB	28334	TCTGCAGACGTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTTTGAT	28275												
QY	121	GATTATAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCACTAGCT	180												
DB	28274	GATTATAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCACTAGCT	28215												
QY	181	GACATACAGCTGCTTTGGTGGAGCTCCAGGATGCTGCTCACTAATAAAGGATACAGCG	240												
DB	28214	GACATACAGCTGCTTTGGTGGAGCTCCAGGATGCTGCTCACTAATAAAGGATACAGCG	28155												
QY	241	GCTACTGATGAGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCGATGCACTT	300												
DB	28154	GCTACTGATGAGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCGATGCACTT	28095												
QY	301	AAAGTTGGCGCGCAAAATTACAGAAATPAGCGAAATATGCTTGGATTAACCAAGCGATTCTT	360												
DB	28094	AAAGTTGGCGCGCAAAATTACAGAAATPAGCGAAATATGCTTGGATTAACCAAGCGATTCTT	28035												
QY	361	GACTCTTTAGGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA	420												
DB	28034	GACTCTTTAGGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA	27975												
QY	421	GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	480												
DB	27974	GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	27915												
QY	481	GGGAAACCGCTGCAATTCCTTAATCTTATGTTGATCAGACATGCTACAGCGACACAG	540												
DB	27914	GGGAAACCGCTGCAATTCCTTAATCTTATGTTGATCAGACATGCTACAGCGACACAG	27855												
QY	541	ATAGAGAAAGATGGAATCGATTAGGATGCAATATTTTGCAGACAGAACGCTAGTGA	600												
DB	27854	ATAGAGAAAGATGGAATCGATTAGGATGCAATATTTTGCAGACAGAACGCTAGTGA	27795												
QY	601	GCTGTAGAAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTACGCTAAAGCAGCA	660												
DB	27794	GCTGTAGAAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTACGCTAAAGCAGCA	27735												
QY	661	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCCAAAT	720												
DB	27734	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCCAAAT	27675												
QY	721	CTTCAAGAACGGNAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAABATATCAACCT	780												

DB	27674	CTTCAAGAACGGNAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATATCAACCT	27615
QY	781	GCAGATGGTCTGTAGTTCAAATCCAGGAACCTACAGTTGGAGCTCAAGCACAGGA	840
DB	27614	GCAGATGGTCTGTAGTTCAAATCCAGGAACCTACAGTTGGAGCTCAAGCACAGGA	27555
QY	841	AGTAGTATGTAGTATTCGTTGTTTCATGCTGTGTAGATGCTGAAAATGAGACCGCT	900
DB	27554	AGTAGTATGTAGTATTCGTTGTTTCATGCTGTGTAGATGCTGAAAATGAGACCGCT	27495
QY	901	TCCATTTTGTGCTGCTGGGTTTCGTGATGATTCACATGTTCAATACGAAAAATCTCTGAT	960
DB	27494	TCCATTTTGTGCTGCTGGGTTTCGTGATGATTCACATGTTCAATACGAAAAATCTCTGAT	27435
QY	961	TCTCAAGTGTCCCAACAGGAGCTCGCAGCACAAAGCTAGACGAGAAAGCGCTGGAGAT	1020
DB	27434	TCTCAAGTGTCCCAACAGGAGCTCGCAGCACAAAGCTAGACGAGAAAGCGCTGGAGAT	27375
QY	1021	GACAGTGTGCTGCGAGCGCTCGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTCTAGGTAAA	1080
DB	27374	GACAGTGTGCTGCGAGCGCTCGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTCTAGGTAAA	27315
QY	1081	GCTGGGCAACCAACAGGCGATCTCAATGCTTTTGAAGACAGATCGCTTCTGCTGCTG- TTGT	1139
DB	27314	GCTGGGCAACCAACAGGCGATCTCAATGCTTTTGAAGACAGATCGCTTCTGCTGCTG- TTGT	27255
QY	1140	GAGCGCAGGAGTTCCTCCGCTGAGCAAGTTCCTATAGGTCATCTGTAAACACAGCTTTA	1199
DB	27254	GAGCGCAGGAGTTCCTCCGCTGAGCAAGTTCCTATAGGTCATCTGTAAACACAGCTTTA	27195
QY	1200	CAAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGC	1259
DB	27194	CAAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGC	27135
QY	1260	TTCAAAATCCATCAATGATGCTATGTTAGGCGCAAGAAATGATGCGATCGTGTGATGAT	1319
DB	27134	TTCAAAATCCATCAATGATGCTATGTTAGGCGCAAGAAATGATGCGATCGTGTGATGAT	27075
QY	1320	AAACATGTTAGTACCCCGCTCTCACAGATCCGTTCTAGAGCAGCAAGAAAGCTCG	1379
DB	27074	AAACATGTTAGTACCCCGCTCTCACAGATCCGTTCTAGAGCAGCAAGAAAGCTCG	27015
QY	1380	AGGACGACAGAAAACACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGAACTCT	1439
DB	27014	AGGACGACAGAAAACACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGAACTCT	26955
QY	1440	TGGAGATGCTATAGTCAAGTTTCGGCACTCAATCTCTTAATGCGAGATCCTCAGTCGAA	1499
DB	26954	TGGAGATGCTATAGTCAAGTTTCGGCACTCAATCTCTTAATGCGAGATCCTCAGTCGAA	26895
QY	1500	TCCTCAAGGGAATATGAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAAGCTCC	1559
DB	26894	TCCTCAAGGGAATATGAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAAGCTCC	26835
QY	1560	ACAGTTTGGCTATCTTATGTCAACTTCTTAATGACTCTACACAGAAAGTTTCATAGCTAA	1619
DB	26834	ACAGTTTGGCTATCTTATGTCAACTTCTTAATGACTCTACACAGAAAGTTTCATAGCTAA	26775
QY	1620	ATTAGAAAGTTTGTGTTGCTGAAGGATCTAGGACAGCAGCTGAATTAAGACGCTTTCCTT	1679
DB	26774	ATTAGAAAGTTTGTGTTGCTGAAGGATCTAGGACAGCAGCTGAATTAAGACGCTTTCCTT	26715
QY	1680	TGAACGAGCTCTTGTGTTTATTCAGCAGGCTGCTCAATATCGCTCTCTATATTCGG	1739
DB	26714	TGAACGAGCTCTTGTGTTTATTCAGCAGGCTGCTCAATATCGCTCTCTATATTCGG	26655
QY	1740	TTATCTCCAAATACCAACACCTAAGTGTGTTTGGAGAGATTATTATGCTTTGGTAAG	1799
DB	26654	TTATCTCCAAATACCAACACCTAAGTGTGTTTGGAGAGATTATTATGCTTTGGTAAG	26595
QY	1800	GCCTTTGTTGAGGCTTACCAAACACTAGAACGATCTTCAATAAATAAAGA	1852

Db 26594 GCCTTTGTTAGGCGCTTACCAACACACTAGAACGATCTTCAATAAATAAAGA 26542

RESULT 8
AX338293
LOCUS AX338293
DEFINITION Sequence 63 from Patent WO0181379.
ACCESSION AX338293
VERSION AX338293.1 GI:18128828
KEYWORDS
SOURCE
ORGANISM Chlamydomophila pneumoniae
Chlamydomophila pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE 1
AUTHORS Bhatia, A., Probst, P. and Stromberg, E.J.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0181379-A 63 01-NOV-2001;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
1..1956
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:8358"
ORIGIN
Query Match 94.4%; Score 1748.8; DB 6; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
Db 205 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264
Qy 61 TCTGACAGCTGGAGTCTAACGACAGCGACCGACCTAGGCTCTCTCCACCGACGCTTGGAT 120
Db 265 TCTGACAGCTGGAGTCTAACGACAGCGACCGACCTAGGCTCTCTCCACCGACGCTTGGAT 324
Qy 121 GATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACTCAACATCACTAGCT 180
Db 325 GATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACTCAACATCACTAGCT 384
Qy 181 GACATACAGGCTGCTTGGTGGAGCTCCAGGATGCTGCTCACTAATAAAGGATACAGCG 240
Db 385 GACATACAGGCTGCTTGGTGGAGCTCCAGGATGCTGCTCACTAATAAAGGATACAGCG 444
Qy 241 GCTACTGATAGGAAACCGGATCGCTCGGAGTGGGAACTAGAGTCCGATGCGAGTT 300
Db 445 GCTACTGATAGGAAACCGGATCGCTCGGAGTGGGAACTAGAGTCCGATGCGAGTT 504
Qy 301 AAAGTTGGCGCGCAATTTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 360
Db 505 AAAGTTGGCGCGCAATTTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 564
Qy 361 GACTCTTTAGTAACTGACTCTTCTTCCGCTTTTACGCTTCTTCTTCTTCTTCTTCTTCTGTA 420
Db 565 GACTCTTTAGTAACTGACTCTTCTTCCGCTTTTACAGGCTGCTTCTTCTTCTTCTTCTTCTGTA 624
Qy 421 GCAACATATAAAGAGCTGAGCTTCTTAAAGAGTGCAGATTAACCCAGTATGCCA 480
Db 625 GCAACATATAAAGAGCTGAGCTTCTTAAAGAGTGCAGATTAACCCAGTATGCCA 684
Qy 481 GGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGACACAG 540
Db 685 GGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGACACAG 744
Qy 541 ATAGAGAAAGATGGAAATGGGATGCAATTTTGGAGGACGACGACGCTAGTGGGA 600
Db 745 ATAGAGAAAGATGGAAATGGGATGCAATTTTGGAGGACGACGACGCTAGTGGGA 804
Qy 601 GCTGTAGAAATGCTTAAATCTAATAACAGTATAAGCAACATAGATTACGCTTAAAGCAGCA 660
Db 805 GCTGTAGAAATGCTTAAATCTAATAACAGTATAAGCAACATAGATTACGCTTAAAGCAGCA 864

661 ATGCTACTGCTTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGGACTCTCCAATT 720
865 ATGCTACTGCTTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGGACTCTCCAATT 924
721 CTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCAAACT 780
925 CTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCAAACT 984
781 GCAGATGGTCTGATGTTCCAAATCAGGAACTACAGTTGGAGGCTCCCAAGCAACAGGA 840
985 GCAGATGGTCTGATGTTCCAAATCAGGAACTACAGTTGGAGGCTCCCAAGCAACAGGA 1044
841 AGTAGATTGCTAGTATTCGTGTTCCATGCTGCTTAGATGCTGCTGAAATAGACGCT 900
1045 AGTAGATTGCTAGTATTCGTGTTCCATGCTGCTTAGATGCTGCTGAAATAGACGCT 1104
901 TCCATTTGATGCTGCGGTTTCCTCAGATGATTCACATGTTCAATACGGAATCTCTGAT 960
1105 TCCATTTGATGCTGCGGTTTCCTCAGATGATTCACATGTTCAATACGGAATCTCTGAT 1164
961 TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACGCAAGAGCGGCTGGAGAT 1020
1165 TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACGCAAGAGCGGCTGGAGAT 1224
1021 GACAGTGTCTGCTGAGCGCTGCGAGATGCTCAGAAAAGCTTTAGAGCGGCTCTAGGTAAA 1080
1225 GACAGTGTCTGCTGAGCGCTGCGAGATGCTCAGAAAAGCTTTAGAGCGGCTCTAGGTAAA 1284
1081 GCTGGGCAACAAAGGCACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
1285 GCTGGGCAACAAAGGCACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1344
1141 AGCGAGGAGTCTCTCCGCTGCGAGCAAGTTCTATAGGGTCATCTGTAACAGCTTTAC 1200
1345 AGCGAGGAGTCTCTCCGCTGCGAGCAAGTTCTATAGGGTCATCTGTAACAGCTTTAC 1404
1201 AAGACCTCAAAATCTCAGGTTCTGATTAATAAACACAGATATCAGCAGCTTTATGCT 1260
1405 AAGACCTCAAAATCTCAGGTTCTGATTAATAAACACAGATATCAGCAGCTTTATGCT 1464
1261 TACAATPCCATCAATGATGCTATGTTAGGCGCGAATGATGCGACTCTGATGTTGATA 1320
1465 TACAATPCCATCAATGATGCTATGTTAGGCGCGAATGATGCGACTCTGATGTTGATA 1524
1321 AACATGTTAGTACCCCGCTCTCAGCAGATCCGTTCTAGAGCAGCAAGAGCTCGA 1380
1525 AACATGTTAGTACCCCGCTCTCAGCAGATCCGTTCTAGAGCAGCAAGAGCTCGA 1584
1381 GGACCAAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGACTCTT 1440
1585 GGACCAAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATAGCAGACTCTT 1644
1441 GAGATGCTCTATAGTCAAGTTTCGGCACTACATCTTAATGAGATCAGCTCAGTGAAT 1500
1645 GAGATGCTCTATAGTCAAGTTTCGGCACTACATCTTAATGAGATCAGCTCAGTGAAT 1704
1501 CCTCAAGCAATATAGGAGATCAGCAAAAGCTTACATCGGAGTGCAGAAAGCTTCCA 1560
1705 CCTCAAGCAATATAGGAGATCAGCAAAAGCTTACATCGGAGTGCAGAAAGCTTCCA 1764
1561 CAGTTTGGCTATCTTATGTTGAACTTTCTAATGACTCTACACAGAGTTCTATAGTAAA 1620
1765 CAGTTTGGCTATCTTATGTTGAACTTTCTAATGACTCTACACAGAGTTCTATAGTAAA 1824
1621 TTAGAAAGTTTGTGCTGAGGATCTAGCAGAGCTGAAATAAAGCACTTCTCTT 1680
1825 TTAGAAAGTTTGTGCTGAGGATCTAGCAGAGCTGAAATAAAGCACTTCTCTT 1884
1681 GAAACCAACTCTCTTGTGTTTATTCAGCAGGCTGCTGCTCAATAATCGGCTCTCTATATCTGCT 1740
1885 GAAACCAACTCTCTTGTGTTTATTCAGCAGGCTGCTGCTCAATAATCGGCTCTCTATATCTGCT 1944

```
QY 1741 TATCTCCAATAA 1752
Db 1945 TATCTCCAATAA 1956

RESULT 9
AX349491
LOCUS AX349491 1956 bp linear PAT 06-FEB-2002
DEFINITION Sequence 14 from Patent WO0202606.
ACCESSION AX349491
VERSION AX349491.1 GI:18615352
KEYWORDS Chlamydomphila pneumoniae
SOURCE Chlamydomphila pneumoniae
ORGANISM Chlamydomphila pneumoniae
REFERENCE 1
AUTHORS Ratti G. and Grandi G.
TITLE Immunisation against Chlamydia pneumoniae
JOURNAL Patent: WO 0202606-A 14 10-JAN-2002;
Chiron S.p.A. (IT)
FEATURES
    Location/Qualifiers
        1..1956
            /organism="Chlamydomphila pneumoniae"
            /mol_type="unassigned DNA"
            /db_xref="taxon:83558"

ORIGIN
Query Match 94.4%; Score 1748.8; DB 6; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
Db 205 ATGAGCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264

QY 61 TCTGCAGAGCTGGACTCAACGACGCGGACCGGACCTTACGCTCCCTCCACCCACGTTTGAT 120
Db 265 TCTGCAGAGCTGGACTCAACGACGCGGACCGGACCTTACGCTCCCTCCACCCACGTTTGAT 324

QY 121 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCACTAGCT 180
Db 325 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCACTAGCT 384

QY 181 GACATACAGCTGCTTTGTGTGAGCTCCAGGATGCTGTCACTAAATATAAGGATACAGCG 240
Db 385 GACATACAGCTGCTTTGTGTGAGCTCCAGGATGCTGTCACTAAATATAAGGATACAGCG 444

QY 241 GCTACTGTATGAGGAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGCAGTT 300
Db 445 GCTACTGTATGAGGAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGCAGTT 504

QY 301 AAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGATTCTT 360
Db 505 AAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGATTCTT 564

QY 361 GACTCTTTAGGTAACTGACTTCTTCGACCTCTTACGCTGCTCTTCTCCAACTCTGA 420
Db 565 GACTCTTTAGGTAACTGACTTCTTCGACCTCTTACGCTGCTCTTCTCCAACTCTGA 624

QY 421 GCATAACATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Db 625 GCATAACATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 584

QY 481 GGGAAAACGCTCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
Db 685 GGGAAAACGCTCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 744

QY 541 ATAGAGAAAGATGGAATGGGATAGGATGCAATTTTGCAGGACAGAACCGTAGTGA 600
Db 745 ATAGAGAAAGATGGAATGGGATAGGATGCAATTTTGCAGGACAGAACCGTAGTGA 804

QY 601 GCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTAAAGCAGCA 660
Db 1741 TATCTCCAATAA 1752
Db 1945 TATCTCCAATAA 1956

805 GCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTTCAGCTAAAGCAGCA 864
861 ATCGCTACTGTAAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAAAT 720
865 ATCGCTACTGTAAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAAAT 924
721 CTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGTCTTAAAAATATCAACCT 780
925 CTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGTCTTAAAAATATCAACCT 984
781 GCAGATGTTCTGTATGTTCCAAATCCAGGAATCAAGTTGGAGGCTCCAAAGCAACAGGA 840
985 GCAGATGTTCTGTATGTTCCAAATCCAGGAATCAAGTTGGAGGCTCCAAAGCAACAGGA 1044
841 AGTAGTATTTGTAGTATTCGTTTCCATGCTGTGTAGATGATGTAAGATGAGACCGCT 900
1045 AGTAGTATTTGTAGTATTCGTTTCCATGCTGTGTAGATGATGTAAGATGAGACCGCT 1104
901 TCCATTTTGTATGCTGGGTTTCGTACAGATGATTCACATGTTCAATACGGAATACTCTGAT 960
1105 TCCATTTTGTATGCTGGGTTTCGTACAGATGATTCACATGTTCAATACGGAATACTCTGAT 1164
961 TCTCAAGCTGCCAACAGGAGCTCGCAGCACAAAGCTAGAGCAGGAAAGCCGCTGGAGAT 1020
1165 TCTCAAGCTGCCAACAGGAGCTCGCAGCACAAAGCTAGAGCAGGAAAGCCGCTGGAGAT 1224
1021 GACAGTGTCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1080
1225 GACAGTGTCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1284
1081 GCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTTTG 1140
1285 GCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTTTG 1344
1141 AGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCTATAGGCTCATCTGTAAAAAGCAGTTTAC 1200
1345 AGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCTATAGGCTCATCTGTAAAAAGCAGTTTAC 1404
1201 AAGACTCAAAATCTACAGGTTCTGATTATATAACACAGATATCAGCAGGTTATGATGCT 1260
1405 AAGACTCAAAATCTACAGGTTCTGATTATATAACACAGATATCAGCAGGTTATGATGCT 1464
1261 TACAAATCCATCAATGATGCTTATGTTAGGCAAGAAATGATCGGACTCTGATGATGATA 1320
1465 TACAAATCCATCAATGATGCTTATGTTAGGCAAGAAATGATCGGACTCTGATGATGATA 1524
1321 AACAAATGAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAAAGCTCGA 1380
1525 AACAAATGAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAAAGCTCGA 1584
1381 GGACAGAAAAAACAAGATCAAGCCCTCGTAGGGTGTATTTCTGGCAATAGCAGAACTCTT 1440
1585 GGACAGAAAAAACAAGATCAAGCCCTCGTAGGGTGTATTTCTGGCAATAGCAGAACTCTT 1644
1441 GGAGATGTTCTATGTTCAAGTTTCGCACTTACAAATCTGTAATGCAGATCACTCAGTCGAT 1500
1645 GGAGATGTTCTATGTTCAAGTTTCGCACTTACAAATCTGTAATGCAGATCACTCAGTCGAT 1704
1501 CTTCAAGCGAAATATGAGGAGATCAGACAAAAGCTTACATCGGAGTGAAGAGCTCGA 1560
1705 CTTCAAGCGAAATATGAGGAGATCAGACAAAAGCTTACATCGGAGTGAAGAGCTCGA 1764
1561 CAGTTTGGTATCTCTTATGTTGCACTTCTAATGATCTTACACAGAACTTCATAGCTAAA 1620
1765 CAGTTTGGTATCTCTTATGTTGCACTTCTAATGATCTTACACAGAACTTCATAGCTAAA 1824
1621 TTAGAAAGTTTCTTGTGCTGAAGGATCTAGGACAGCTGAAATAAAAAGCATTTCCTTT 1680
1825 TTAGAAAGTTTCTTGTGCTGAAGGATCTAGGACAGCTGAAATAAAAAGCATTTCCTTT 1884
1681 GAAACGAACTCCTGTTTATTCAGCAGGTGCTGTTCAATATCGGCTCTCTATATTTCTG 1740
1885 GAAACGAACTCCTGTTTATTCAGCAGGTGCTGTTCAATATCGGCTCTCTATATTTCTG 1944
```

[illegible]

Qy	121	GATTATTAAGACTCAAGCGCAAAACAGCTTACGATACATCTTTACTCTCAACATCACTAGCT	180
Db	1093	GATTATTAAGACTCAAGCGCAAAACAGCTTACGATACATCTTTACTCTCAACATCACTAGCT	1152
Qy	181	GACATACAGGCTGCTTTTGGTGAGCCTCCAGGATCTCTCACTAATAATAAGGATACAGC	240
Db	1153	GACATACAGGCTGCTTTTGGTGAGCCTCCAGGATCTCTCACTAATAATAAGGATACAGC	1212

ORGANISM	Chlamydia pneumoniae	Db	1510	ATAGAGAAAGATGAAATCGATTAGGGATGCAATATTTTCAGGACAGAACGCTAGTGA	1569
REFERENCE	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia	Qy	601	GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCTAGTAAAGCAGCA	660
AUTHORS	Murdin, A.D., Omen, R.P., Wang, J. and Dunn, P.	Db	1570	GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCTAGTAAAGCAGCA	1629
TITLE	i (Chlamydia) antigens and corresponding dna fragments and uses	Qy	661	ATCGCTACTGTCTAAGACACAAATAGCTGAGGCTCAGAAAAGTTCCCGGCTCTCCAAAT	720
JOURNAL	Therof	Db	1630	ATCGCTACTGTCTAAGACACAAATAGCTGAGGCTCAGAAAAGTTCCCGGCTCTCCAAAT	1689
FEATURES	Patent: WO 006739-A 5 09-NOV-2000;	Qy	721	CTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATATCAAACT	780
source	Location/Qualifiers	Db	1690	CTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATATCAAACT	1749
	1. 2238	Qy	781	GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACAGGA	840
	/organism="Chlamydia pneumoniae"	Db	1750	GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACAGGA	1809
	/mol_type="unassigned DNA"	Qy	841	AGTAGTATTTGGTAGTATTCGTGTTTCATGCTGCTGTAGATGATGCTGAAATGAGCCGCT	900
	/db_xref="taxon:83558"	Db	1810	AGTAGTATTTGGTAGTATTCGTGTTTCATGCTGCTGTAGATGATGCTGAAATGAGCCGCT	1869
	/note="unnamed protein product"	Qy	901	TCCATTTTGTATGCTCGGTTTCCTCAGATGANTTCAATGTCGGAATAATCTCTGAT	960
	/transl_table=1	Db	1870	TCCATTTTGTATGCTCGGTTTCCTCAGATGANTTCAATGTCGGAATAATCTCTGAT	1929
	/protein_id="CAC17330.1"	Qy	961	TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAAGCTAGACGCAAGAACGCCGTGGAGAT	1020
	/db_xref="GI:1134373"	Db	1930	TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAAGCTAGACGCAAGAACGCCGTGGAGAT	1989
	/translation="MVNPIGPGPIDEETERTPPADLSAQGLEASANKSABQRIAGAE	Qy	1021	GACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA	1080
	AKPKESKTSVERMSILRSNAVLADKLGIASSNSSSSTSRSDVDTATATP	Db	1990	GACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA	2049
	PPPTFDYKTAQATYDTFTSTSLADIQAALVSLQDVTNIXDKTADTETAAPW	Qy	1081	GCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG	1140
	ETKNADAVKGAQITELAKYASDNQAILDSLGKLTSPDLQALLOSVAANNKAEEL	Db	2050	GCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG	2109
	KEMQNPVPGKTPAIQSGVDQTDATQIEKGNARIDRIFAGQWASGAVENAKSN	Qy	1141	AGCGCAGGAGTCTCTCCGCTGCAGCAAGTTCTATAG 1177	
	NSINIDSAKAATATKQVIAAOKFPDPSILOEABOMVIOAEKOLKNTKPADGSDV	Db	2110	AGCGCAGGAG-TACTCCGCTGCAGCAAGTTCTATAG 2145	
	PNPTGTVGGKQSSSISIRVSLMLDADENETASILMSGFRQIMHFNENTPDSQAA				
	QOELAAQARAAKAGDSDAAALADAKALEALGKAGQOQILNALGQIASAVVSA				
	GVLFQVLMIRARYQAYVEQLISEEDLNSAVDHHHH"				
ORIGIN					
Query Match	62.7%; Score 1161.8; DB 6; Length 2238;				
Best Local Similarity	99.7%; Pred. NO. 3e-265;				
Matches 1174; Conservative	0; Mismatches 2; Indels 1; Gaps 1;				
Qy	1	ATGAGTCTGGCAGATAAGCTGGTATTCCTCTAGTAACAGCTCGTCTTCTACTAGCAGA	60		
Db	970	ATGAGTCTGGCAGATAAGCTGGTATTCCTCTAGTAACAGCTCGTCTTCTACTAGCAGA	1029		
Qy	61	TCTGACAGCTGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCACGCTTGAT	120		
Db	1030	TCTGACAGCTGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCACGCTTGAT	1089		
Qy	121	GATTATAGACTAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACTAGCT	180		
Db	1090	GATTATAGACTAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACTAGCT	1149		
Qy	181	GACATACAGGCTGTTTGGTGAGCTCCAGGATGCTGTCATTAATATAAGAGTACAGCG	240		
Db	1150	GACATACAGGCTGTTTGGTGAGCTCCAGGATGCTGTCATTAATATAAGAGTACAGCG	1209		
Qy	241	GCTACTGATGAGAAACCGCAATCGCTCGGAGTGGGAACTAAGATGCCGATGCAATT	300		
Db	1210	GCTACTGATGAGAAACCGCAATCGCTCGGAGTGGGAACTAAGATGCCGATGCAATT	1269		
Qy	301	AAAGTTGCGCGCAATACAGATACGGAATATGCTTCGGATAACCAAGCATCTTT	360		
Db	1270	AAAGTTGCGCGCAATACAGATACGGAATATGCTTCGGATAACCAAGCATCTTT	1329		
Qy	361	GACTCTTTTAGTAACTGACTCTCTTTCGACCTCTTACAGGCTGCTTCTCTCAATCTGTA	420		
Db	1330	GACTCTTTTAGTAACTGACTCTCTTTCGACCTCTTACAGGCTGCTTCTCTCAATCTGTA	1389		
Qy	421	GCACAAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	480		
Db	1390	GCACAAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	1449		
Qy	481	GGGAAAACGCTGCAATCTTCTTGTATGATCAGACAGATGCTTACAGGACACAG	540		
Db	1450	GGGAAAACGCTGCAATCTTCTTGTATGATCAGACAGATGCTTACAGGACACAG	1509		
Qy	541	ATAGAGAAAGATGGAATGCGATAGGATGCATATTTTTCAGGACAGAACGCTAGTGA	600		

RESULT 13

AX045135

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

1. 1456

/organism="Chlamydia pneumoniae"

/mol_type="unassigned DNA"

/db_xref="taxon:83558"

101. >1452

/note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CAC17379.1"

/db_xref="GI:11343735"

/db_xref="RENTREMBL:CAC17379"

/translation="MVNPIGPGPIDEETERTPPADLSAQGLEASANKSABQRIAGAE

AKPKESKTSVERMSILRSNAVLADKLGIASSNSSSSTSRSDVDTATATP

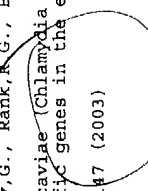
PPPTDDYKTAQATYDTFTSTSLADIQAALVLSQDQAVTNKIDTAANTDEBTAAAEW
ETKNADAVKGAQITELAKYASDNQAIILDSGLKLSFDLLOALIQSVANNKKAELL
KEMONDVPVPGTAPTAQSLVDOTATATQTEKQGNARIDAYFACNAGSAGVNAKSN
NSINISDAKAAITAKTQILBAOKKFPDPSILOEAKOMVIOAEKDLKNIKPADGSDV
PNPPTVGGSKQSSSISIRVSMLLDAAENETASILMSGFRQMHMENTENPDSQAA
QQLAQAFAAFAAAGGDSAAALADAQKALENNAKAGCQSQSILNAGLQIASAAVVA
G"

ORIGIN

Query Match: 62.1%; Score 1151; DB 6; Length 1456;
Best Local Similarity 100.0%; Pred. No. 1,1e-262;
Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
Db 305 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 364
QY 61 TCTGACAGCTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTTTGTAT 120
Db 365 TCTGACAGCTGGACTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTTTGTAT 424
QY 121 GATTATAAGACTCAACGCAAAACAGCTTACGATACCTATCTTTACCTCAACATCACTAGCT 180
Db 425 GATTATAAGACTCAACGCAAAACAGCTTACGATACCTATCTTTACCTCAACATCACTAGCT 484
QY 181 GACATACAGGCTGCTTTGGTGAGCCCTCCAGGATGCTGTCACTAATAAAGGATACAGCG 240
Db 485 GACATACAGGCTGCTTTGGTGAGCCCTCCAGGATGCTGTCACTAATAAAGGATACAGCG 544
QY 241 GCTACTGATGAGGAACCCCAATCGCTGGGAGTGGGAACCTAAGAACTCCGATCGAGTT 300
Db 545 GCTACTGATGAGGAACCCCAATCGCTGGGAGTGGGAACCTAAGAACTCCGATCGAGTT 604
QY 301 AAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGATTCT 360
Db 605 AAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGATTCT 664
QY 361 GACTCTTTAGGTAAATGACTCTTCCTTCGACCTCTTACAGCTGCTCTTCCCAATCTGTA 420
Db 665 GACTCTTTAGGTAAATGACTCTTCCTTCGACCTCTTACAGCTGCTCTTCCCAATCTGTA 724
QY 421 GCAACAATAACAAACGACGCTGCTCTTAAAGAGATCGAATACCAAGTACCGCTAGTCCCA 480
Db 725 GCAACAATAACAAACGACGCTGCTCTTAAAGAGATCGAATACCAAGTACCGCTAGTCCCA 784
QY 481 GGGAAACCGCTGCAATGCTCAATCTTGTAGTTGATCAGACAGATGCTACAGCGACACAG 540
Db 785 GGGAAACCGCTGCAATGCTCAATCTTGTAGTTGATCAGACAGATGCTACAGCGACACAG 844
QY 541 ATAGAGAAAGATGGAATGGATTAGGATGCAATTTTGCAGGACAGACGCTAGTGA 600
Db 845 ATAGAGAAAGATGGAATGGATTAGGATGCAATTTTGCAGGACAGACGCTAGTGA 904
QY 601 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACGCTAAAGCAGCA 660
Db 905 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACGCTAAAGCAGCA 964
QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTCCCGACCTCCCAAT 720
Db 965 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTCCCGACCTCCCAAT 1024
QY 721 CTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCAACCT 780
Db 1025 CTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCAACCT 1084
QY 781 GCAGATGGTTCGTGATGTTCCAAATCAGGAACCTACAGTTGGAGGCTCCCAAGCAACAGGA 840
Db 1085 GCAGATGGTTCGTGATGTTCCAAATCAGGAACCTACAGTTGGAGGCTCCCAAGCAACAGGA 1144
QY 841 AGTAGTATTGGTAGTATTCGTGTTTCCATGCTCTTGTAGATGATGCTGAAATGACCGCT 900
Db 1145 AGTAGTATTGGTAGTATTCGTGTTTCCATGCTCTTGTAGATGATGCTGAAATGACCGCT 1204

QY 901 TCCATTTTGATGTCGTGGTTCGTTCAGATGATTCATGTTCAATACGGAATAATCCTGAT 960
Db 1205 TCCATTTTGATGTCGTGGTTCGTTCAGATGATTCATGTTCAATACGGAATAATCCTGAT 1264
QY 961 TCTCAAGCTGCCCAACACAGGAGCTCGCAGCACAAAGCTAGCAGCAGCGAAGCCGCTGAGAT 1020
Db 1265 TCTCAAGCTGCCCAACACAGGAGCTCGCAGCACAAAGCTAGCAGCAGCGAAGCCGCTGAGAT 1324
QY 1021 GACAGTGTCTGTGACGCGCTCGCAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGGTAAA 1080
Db 1325 GACAGTGTCTGTGACGCGCTCGCAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGGTAAA 1384
QY 1081 GCTGGGCAACACAGGCGATCTCAATGCTTTAGCAGACAGCTTCTGCTGCTCTTGTG 1140
Db 1385 GCTGGGCAACACAGGCGATCTCAATGCTTTAGCAGACAGCTTCTGCTGCTCTTGTG 1444
QY 1141 AGCGCAGGAGT 1151
Db 1445 AGCGCAGGAGT 1455

RESULT 14
AE016994
LOCUS Chlamydomophila caviae GPIC section 1 of 4 of the complete genome. 304769 bp DNA linear BCT 15-AUG-2003
DEFINITION Chlamydomophila caviae GPIC
ACCESSION AE016994 AE015925
VERSION AE016994.1 GI:29834116
KEYWORDS
SOURCE Chlamydomophila caviae GPIC
ORGANISM Chlamydomophila caviae GPIC
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia. 
REFERENCE 1 (bases 1 to 304769)
Read, T.D., Myers, G.S., Brunham, R.C., Nelson, M.C., Paulsen, I.T., Heidelberg, J., Holtzaple, E., Khouri, H., Fedorova, N.B., Carty, H.A., Umayam, L.A., Haft, D.H., Peterson, J., Beanan, M.J., White, O., Salzberg, S.L., Hsia, R.C., McClarty, G., Rank, M.G., Bavoil, P.M. and Fraser, C.M.
Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae
Nucleic Acids Res. 31 (8), 2134-2147 (2003)
JOURNAL
MEDLINE 22569155
PUBMED 12682364
REFERENCE 2 (bases 1 to 304769)
Read, T., Myers, G., Brunham, R., Nelson, M., Paulsen, I., Heidelberg, J., Holtzaple, E., Khouri, H., Fedorova, N.B., Carty, H., Umayam, L., Haft, D., Peterson, J., Beanan, M., White, O., Salzberg, S., Hsia, R., McClarty, G., Rank, R., Bavoil, P. and Fraser, C.
Direct Submission
Submitted (29-OCT-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1..304769
/organism="Chlamydomophila caviae GPIC"
/mol_type="genomic DNA"
/isolate="GPIC"
/db_xref="taxon:227941"
57..1052
/gene="hemB"
/locus_tag="CCA00001"
57..1052
/gene="hemB"
/locus_tag="CCA00001"
/note="similar to GB:U16784, SP:P33395, and PID:290384; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="delta-aminolevulinic acid dehydratase"
/protein_id="AAP04754.1"
/db_xref="GI:29834117"
translation="MSSLLALLRRPRNRKTAARLVAETALLPDFICPFVKEGKN
ICEETESLPGVVRSIDILLKEISLGLKAVILFPVIPSIDLDKDAYSSNPKN
LCKSIYEVKKAAPPDLCLVTSIDALDPYTHGHGIMDQGEVLNDSVRFIGNIATLHAE


```

MGADIVAPSDMGRVAHIRAKLDQSGWTNTLILSVSKYASSLYNPFPRDALGSHLQS
GDRNYQMPKNVLEALCESLDEQEGADMLMKPAGLYDLVLRVNSSTTLPLAAQY
VSGEYALIAAADKMGMLNKEAMFYSLIAIKRAGADMIIISYATPLILEMLVSKF"
complement(1068. .2480)
/gene="ngra"
/locus_tag="CCA00002"
complement(1068. .2480)
/gene="ngra"
/locus_tag="CCA00002"
note="similar to SP:Q56586, GB:L16784, SP:P33395, and
PID:290384; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, subunit A"
/protein_id="AAP04755.1"
/db_xref="GI:29834118"
/transl_table=11
translation="MKIAIRGRLDLSQSPKESGFLKRIIDPALVSDLRPYSAALAK
LKYEDDIVSSGPIAEYKPNPGVFIITPVSGIVKEIRGKEKSLDVLVTKPGQNL
TEYSYDLKSIAELLEIFKEGFLFKORPDIPALPTQTRDVFNIADNRPTTP
STKQIAVSSREGEFTVNVGRVAVAKFLGCPHIVSTDRALPKDLKSIHLHKI
TGYPGSPSTHIYAPIITNEXDIVFTLSFQEVLTGHLFKGRILNEQVVALAGSG
LKSRLRYVITTKGANFESLLPLQIDSSVLSLIGDPLTGLCDRESLPLCLGNRDSIT
SVLPNPKROAFNFGINKPHTIRTYLGSFLKRGKHYMDPOTNLHGETRPIIDTEI
YDKVMPKIPVPLIKAVITKDFELACMLGLEVCPEDFALPTFIDPSKTEMLTIKID
ALKDYAKETGILNLEYDKS"
complement(2502. .2936)
/locus_tag="CCA00003"
complement(2502. .2936)
/locus_tag="CCA00003"
note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP04756.1"
/db_xref="GI:29834119"
/transl_table=11
translation="MNSNSAKIIDSIKQILTLVNIIDPSPGSSLSDDSEVDVEYIM
EKTEKIQELDKSDILQOTQMTQOMEVFNPNPNFSPPEMLALENIRSSCDAYKK
ETALINEVTKDUGGETKTSNNKTKTSSKKNKKKNWPL"
3048. .5201
/gene="greA"
/locus_tag="CCA00004"
3048. .5201
/gene="greA"
/locus_tag="CCA00004"
note="similar to GB:L10328, SP:P04846, GB:M12163,
PID:146649, PID:290510, GB:U00096, and PID:1790093;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="transcription elongation factor GreA"
/protein_id="AAP04757.1"
/db_xref="GI:29834120"
/transl_table=11
translation="MDYLEKIQVLIDEEQSPSFNLWEYCFNDVVRGAELVQILEKV
KHSLAPLFQKIRIVLPLWEMIPDGKEDRVLOLVLDQVNTAKPYDAIDYVNRK
YAGRENEALRVGLDRGEFOYLSRDFPLMLHSCGNFVHSGVGVGEVSVSL
QKVLIFEGGVMAKDISFTAFKSLIPLKDHFLSRPDGDPGFESVAKEHSEVIE
MLLKDLGPKAKEIKDELVDLVIPEAEWRWQATKIKKNCITITPKIKIYYRXN
NQSLEIQTRLSQEDTSKLVRIYQIRDSHSLKPNREIVIKQALQILSVRG
NQSLIQRLDLSLDFLGESVLDSEFLSPEEIVISVINISIVVLQAFMLIKK
HFWPNVNFKIFLSTSPSLRELTFKVFREGDSYRKIKKLECAEHMPTPEVFT
WFFLKGSHDDGIEPKNDKIERLFLEALVFMVHASTPQKELGKKIYFVLSQRYL
ARNVMEIGAPLYLKEILLSTKCSQFSSDLSVLQSLAEVHPDLKNTVVEEDIL
WTPESFTKWNKLOSGKEMVDNAIEEDARVLGDLRENSKFALEKRAKRLQEI
RVLSSEVNRARILTKDIVFTDKVGCGKSLDENGMLTISYILGPDWADPDNFIJSL
KSKLAQEMLDKSGAAVQFGKKYKISRIQSIWDE"
5353. .6543
/locus_tag="CCA00005"
5353. .6543
/locus_tag="CCA00005"
note="similar to SP:P04693, GB:L10328, SP:P04846,
GB:M12163, PID:146649, PID:290510, GB:U00096, and
PID:1790093; identified by sequence similarity; putative"
/codon_start=1

```

```

/transl_table=11
/product="amino acid biosynthesis aminotransferase, class
I"
/protein_id="AAP04758.1"
/db_xref="GI:29834121"
/transl_table=11
translation="MSFFNQLTFFSPDSILGLQKLFLEDERBEKYNLVIGVYEDPSKA
YGGFSVRAQFLFLEDENNKGLFISGLSSENQENMEKLVFGDVNSFVUGAALCGT
GALHIGAKIFSMAYTSGKYVIEPQTWGNHVRIFAQOGLVLYKYPIYSSSKSLFDEM
LSVLKAAKPNLSVLQCCCHNPTGMDLNEDMVRITAEVNMKEHQLFFPTFYILGFRG
IKEDRRPIEIFTEVGNPVFAACAKNFVGYFAVHSKIVTEIDLDKIISCLCEK
ARGEYSSPRHGAKIVTITLSDASLKKEWLSLDTIRSSIGKTRARLVAMRNHVCHS
PFILLQKGFPGYFGFSEVQVFLRSEKGIYITSGARFNLNGITDKNIDHVQSFAEA
YQLS"
6839. .7861
/locus_tag="CCA00006"
6839. .7861
/locus_tag="CCA00006"
note="similar to GB:L10328, SP:P04846, GB:M12163,
PID:146649, PID:290510, GB:U00096, and PID:1790093;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP04759.1"
/db_xref="GI:29834122"
/transl_table=11
translation="MAGSVAYRRHRKNRNVYLVIALSVILCWSLPRPYENIQKHVF
SYTRVFFKSGVPSIDYTIQDVENIILKDRIALIEERLQAEIANYTPPVSEILSP
YFKRTISRVYRDFAHMGSSCWINNKQKIQKNSPVLSGVLGVGLVYVGGQSRRI
RLITDVGMPQSVIAGVGIQAWLVQDIOQLTKHLERLSDYLEKDKYQVQDLBEL
SSSIOGSDNTLLRGTLSSGSGPLMKDFTLVHGBGFCFSDGKLRKVDGILITGLD
GVYFPGFLVAEITKVQCPRGACSYKLEAKSLAEDLTNLSSVLLPAMEFNPDRDPI
FGLIWD"
complement(7821. .10958)
/gene="recB"
/locus_tag="CCA00007"
complement(7821. .10958)
/gene="recB"
/locus_tag="CCA00007"
note="identified by match to TIGR protein family HMM
TIGR01866"
/codon_start=1
/transl_table=11
/product="exodeoxyribonuclease V, beta subunit"
/protein_id="AAP04760.1"
/db_xref="GI:29834123"
/transl_table=11
translation="MKPFDIFNPQTSIRKQYFLEASAGTKTTFIBQIVLRALLEGSV
SHVENLAVFTNATNELKRIQNLKQALQIKSVLEDPSPSKLPPYLKQTSNVKLL
YQOVNALATIDRMALFTTHGFCNVVLQGHFPMQMTQKNAALTHSQAVFHHIKYLS
QDLWRSVLPFOFYLLAAKYNNSKHTAFLDKLSSYTLOTPEHLQPTSTTLDALRS
MNSIRSKQIDIPKEDFLQILQHKDGFKKQPPIDEDLRLFVLCVASETSVQLPSF
SKIAETPHKRLTRYKSKASHIEETLWSLYTEQFCNIELFNTLLQDQVLIYKXH
YTOWLSPDSIIALEIILSSREEIETLRKRFQVLIDEFQDTRKQNNIFSKLFA
NDAFSGSLFIDGPQKSIYWRNADIATYLKAKSPKSSQLHNNYSITQALMEAV
NIFLCKSPFLQTPSYEPLCPQSSHEFNSHAPTHFFSDYDIDQAAWISHT
ASYLQATYVILPLGRMALLVSDAQAPDLTHGSIIPVAFKSKSIPIHLTLYTLTAWL
EAILHPENYKIQRLVLSLFRHNSLDTTEKKELYSTVFPSRVSVDYVGLLATVFFH
MILQBALLKTPQGDLSFQEMERLCALYTVSSHPOQLLYLQYFSETGRWESLSFS
SVSEDEILKIIITTHASKGLEVDVFCPGDKSKKNSASEMIRENIVCTRAKKQLF
IPQNSNSGRONALTNYKHYGVSIEDLAQLTKEHRLPFSITRSETDTPV
YKIIPTSTLTAPAKQIFSFSSVKLAFDNESVDNVAEPVASFVLPGRKATGII
HKILNISPKNFISKSIIVTHFVNKTHLEGYEEIIQKLVTFSSFLSPSATESFS
LQDICPSKACRESFLPSNQEQOLWQGVIDLFEYNDKYIIDWKTSLGETSLSEYKE
NLLNYIKEQNLQYQGAITYHAARKFQQPDIIVSDVEMGFIRGMDPEGNGLCLPQ
KISNFTTRKCFAYH"
complement(10945. .14010)
/gene="recC"
/locus_tag="CCA00008"
complement(10945. .14010)
/gene="recC"
/locus_tag="CCA00008"
note="similar to GB:L25079, SP:P42822, and PID:529423;
identified by sequence similarity; putative"
/codon_start=1

```



```

/transl_table=11
/product="exodeoxyribonuclease V, gamma subunit"
/protein_id="AAP04761.1"
/db_xref="GI:29834124"

Query Match
Best Local Similarity 53.8%; Score 361.2; DB 1; Length 304769;
Matches 903; Conservative 0; Mismatches 733; Indels 42; Gaps 6;

Qy 104 CTCACCCACGCTTTGATGATTTATAGCTCAAGCGCAACAGCTTACGATCTACTTTA 163
Db 21975 CTCCTCAAACTTTTATGATGAAGCTAAGACGCAAGCAGAGATGCGTTCGAGA 22034
Qy 164 CCTCAACATCACTAGCTACACATACAGGCTGCTTTGGTGAGCTCCAGAGATCTGTCACTA 223
Db 22035 GTGCGACTTATGATCAGTTCAAGACCGCTTTACAGCAGCTGCAAGATCTGTGAAC 22094
Qy 224 ATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAACTA 283
Db 22095 AGATGGAGCAATTAGCTACTACTGATGCAGAAAAGCTTACAGTTTGCTACATGGAAAACGG 22154
Qy 284 AGAATGCCGATGCGATTAAGTTGGCGGCAAAATTACAGAAATTAGCGAAATATGCTTCGG 343
Db 22155 CTCCTTGAGCGGAGAAAGTACGCTGGGATACACTTAAACCGTTGGTGTCTATTTCTTAAG 22214
Qy 344 ATAAACCAAGCGATTTCTTGACTCTTTAGGTAACTGACTTTCCTTCGACCTCTTACAGGCTG 403
Db 22215 AGNACCAGAAGCTTCTTGAGGCAATAAAGACGACCTCGTCTATGGATCAGATTATGGGAG 22274
Qy 404 CTCCTTCCAAATCTGTAGCAACAATTAACAAGCAGCTGAGCTTCTTTAAAGAGATGGAAG 463
Db 22275 CTGCGGGAACAAGTAGAACCATAAACAACCTGCTGAGAGGTTAATTAACAGAGTTGAAGG 22334
Qy 464 ATAACCCAGTAGTCCCGGGAACCGCTGCAATTTGCTCAATTTAGTTGATCAGACAG 523
Db 22335 A-----AGCTGGGTTAGCTATCTCTGTGATGATGACCTTGAGAAAGCAATTA 22382
Qy 524 ATGCTACAGCAGACAGATAGAGAAAGATGGAATTCGGATTAGGATGCATATTTTGAG 583
Db 22383 CAACCTCAGGAACCTCAGGTTACTGAAATTAGCAGATGCTATATCGGAAGCTTATGCTCGG 22442
Qy 584 GACAGACGCTAGTGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATGACAACTAG 643
Db 22443 GGAATAACAGTACCGCGGCTGTGGGGAAGCAGCAGGCAATAACAGCCCGCAATATAG 22502
Qy 644 ATTCAAGTAAAGCAGCAATCGTACTGTCTAGACACAAATAGCTCAAGCTCAGAAAAAGT 703
Db 22503 AAGCTTCCAAACAAACTATTGCAATGCAAAAAAGTCAAGAACGCTCTTAACCTTG 22562
Qy 704 TCCCGACTCTCCAAATCTTCAAGAACGCGGAACAAATGGTAAATACAGCTGAGAAAGATC 763
Db 22563 CTCAGATTTCCGATACTCAAAAGCTGTTTGAAGAACAACAACAGCAGCAGCAAAAGATA 22622
Qy 764 TTAATAATCAAACTGCAGATGCTTCTGATGTTCCAAATCCAGAAC-----812
Db 22623 TCCTCAATGTGAACCTAGTGTGTAGTGTGCTGCTATCGGTGCTCGGAGCTCTCTG 22682
Qy 813 -TACAGTTGGAGGCTCCAGCAACAAAGGAAGTAGTATTGGTAGTATTCGTGTTCCATGC 871
Db 22683 GTAGTGTGGGACTTCTCAAAATCGCGGTGCTACTTAGGGGAAGTTGCGTATCGATGT 22742
Qy 872 TGTTAGATGATCTGAAATAGACCGCTTCCATTTTGATGTCTGGTTTCGTCTAGATGA 931
Db 22743 TATTGACTGATGTTGATATGAACCGCAGCGATCATTTATGCAAGTTTTCAGAAATATGA 22802
Qy 932 TTCATATGTTCAATACGGAATTCCTGATTTCAAGCTGCCCAACAGGAGCTCGCAGCAC 991
Db 22803 TCGATAAATTCCTATGATCAAAATCTCTGATTTTACAGCGCTTTTAGAAGAGATTTATGAATC 22862
Qy 992 AAGCTTA-----GAGCAGCGGAAGCGCTGGAGATCACAGTGTCTGCTGAGCGCTGG 1042
Db 22863 AAGTAAACCGACTTATCAACGCGAGATCAATCTCTGCAAGTCCGGAAGCTACAGCAACTAC 22922

```

```

Qy 1043 CAGATGCTCAGAAAGCTTTAGAAGCGGCTCTAGGTAAGCTGGSCAACAAACAGGGCATAC 1102
Db 22923 AAGAAATACAAACAAACCATACAAGTGCCTTCAAGGGAATCGCGGTCAAGACGCGCATGA 22982
Qy 1103 TCAATGCTTTAGGACAGATCGCTTCTGCTCTGTTGTGAGCGGAGGAGTTCCTCCCGCTG 1162
Db 22983 TCAATGCTTTAGGAGCTTATAACAAACAGCAGCTTCAATTTCTACAGGAGCTCCTATCGCTT 23042
Qy 1163 CAGCAAGCTTCTATAGGCTCATCTGTAAACAGCTTTTACAAGAC---CTCAAAATCTACAG 1219
Db 23043 CTGCAATCAAGGTGGATCAGCTGTAAGCAGCTTTTACAAACAGGATCTACTCTGCGA 23102
Qy 1220 GTTCTGATATPAAACAACA---GATATCAGCAGGTTATGATGCTTACAATCCATCATG 1276
Db 23103 GTTCTAAATCTTACGCGGATTCCTTATCTGCGAGGTATGGGCGCATCAATCTTTTAAATG 23162
Qy 1277 ATGCTATGTTAGGCGCAGAAATGATCGACTCTGTGATGTGATAAACAATGTAAGTACCC 1336
Db 23163 ATGTGTTACTACG---TAGTAGTGCACTTAACCGTGAGTTTATAGTCGTACATCGATC 23219
Qy 1337 CCGCTCTCACACGATCCGTTCTTAGAGCAGAAACAGAAAGCTCGAGGACCGAGAAAAACAG 1396
Db 23220 CAGCATTAACGACAGACAGTTTCTAGAACAGAAACTCGGCTCGTGATAATGATAACGCG 23279
Qy 1397 ATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTTGGAGATGTTCTATAGTC 1456
Db 23280 CTCAGCGTTTTCGAAGAACTATAGCTGCTAATAGTAATCTTCTGGGATGTTTATGCAAT 23339
Qy 1457 AAGTTTCGGCACTCAATCTGTAATGAGATCACTCAGTCGAATCCTCAAGCGAATAATG 1516
Db 23340 CCGTAGGTGTTATGCAAACTGCTAGTGTATTAACAATAATCCCAACAGCGAATGAG 23399
Qy 1517 AGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTCCACAGTTTCGGCTATCCTT 1576
Db 23400 AAGAAATCAACACAGAGCTCACTCTGAGGTTACGAAAGCTCCGAGCTCAGGTTATCCTC 23459
Qy 1577 ATGTGCAACTTTCTAATGACTCTACAGAGTTTCATAGCTTAATAGAAAGTTTGTGTTG 1636
Db 23460 ATGTACAGCTTTCTTAACAGCTCTACGAAAGAGTTCAATGCTCAACTCGAAGTGAATTTG 23519
Qy 1637 CTGAAGGATCTPAGNACAGCAGCTGAAATTAAGCACTTTCCTTTGAAACGAACTCCTTGT 1696
Db 23520 TTCAGGATTCGAAAGACTTCCCGAAGCAAAAGAGCTGCGTTTGAGAAACAGCCTTGT 23579
Qy 1697 TTATTCAGCAGGCTCGTCAATATCGCTCTCTATATTCGTGTTATCTTCAATAACA 1754
Db 23580 TCATCCAGCAGGTATTAGTGAACGATGATCTCTGTTCTCGGATACCTACAGTAAAA 23637

```

```

RESULT 15
BD264444
LOCUS
DEFINITION
    Compounds and methods for treatment and diagnosis of chlamydial
    infection.
ACCESSION
    BD264444
VERSION
    BD264444.1 GI:33074212
KEYWORDS
    JP 2002531129-A/32.
SOURCE
    Chlamydia trachomatis
ORGANISM
    Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE
    1 (bases 1 to 1537)
AUTHORS
    Probst,P., Bhatia,A., Skeiky,Y.A.W., Fling,S.P., Jen,S. and
    Stromberg,B.J.
TITLE
    Compounds and methods for treatment and diagnosis of chlamydial
    infection
JOURNAL
    Patent: JP 2002531129-A 32 24-SEP-2002;
CORIXA CORP
OS
    Chlamydia trachomatis
PN
    JP 2002531129-A/32
PD
    24-SEP-2002
PF
    08-DEC-1999 JP 2000586916
PR
    08-DEC-1998 US 09/208277, 08-APR-1999 US 09/288594 PR
    01-OCT-1999 US 09/410568, 22-OCT-1999 US 09/426571 PI PETER

```

```

PROBST,AJAY BHATIA,YASIR A W SKEIKY,STEVEN P FLING,PI SHYIAN
PI, ERICA JEAN STROMBERG
PC C12N15/09,C12N15/09,A61K35/12,A61K38/00,A61K39/118,A61K39/395,
PC A61K39/395,
PC A61K45/00,A61K48/00,A61P33/00,C07K14/295,C07K16/12,C07K19/00,
PC C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/PC
532,
PC
G01N33/569//C12N1/19,C12R1/01,C12N15/09,C12R1/01,C12Q1/68,PC
C12R1/01,
PC C12N15/00,C12N15/00,A61K37/02,C12N15/00,C12R1/01)CC
Compounds and methods for treatment and diagnosis of CC
chlamydial infection
FH key Location/Qualifiers
FT source 1..1537
FT 1..1537 /organism='Chlamydia trachomatis'.
FEATURES
source
1..1537
Location/Qualifiers
/organism='Chlamydia trachomatis'
/mol_type='genomic DNA'
/db_xref='taxon:813'
ORIGIN
Query Match 10.1%; Score 187.2; DB 6; Length 1537;
Best Local Similarity 53.0%; Pred. No. 1.6e-33;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY 863 TTTCATGCTGTAGATGCTGAAATGAGACCGCTTCCATTTTGGATCTGGGTTTC 922
Db |||||
186 TTTCCTTGTGCTTGTAGATGATGACAAATGAAATGGCAGCGAATGCAATGCAAGGTTTC 245
QY 923 GTCAGATGATTCATGTTCAATACGAAATCTGATTTCTCAAGCTGCCACAGAGC 982
Db |||||
246 GATCTATGATCGAACAAATTAATGTAACAAATCTGCAACAGCTAAAGAGCTACAAGCTA 305
QY 983 TCGCAGCACAAAGCTAGACGACGAAAGCCGCTGGAGATGACAGTGTCTGTCAGCGCTGG 1042
Db |||||
306 TGGAGGCTCAGCTGA---CTGCGATGTCAGATCAACTGGTTGGTGGGATGGCGAGCTCC 362
QY 1043 CAGATGCTCAAGAGCTTTAGAGCGGCTCTAGTAAAGCTGGGCAACAAC----- 1093
Db |||||
363 CAGCCGAATACAGCAATCAAGATGCTCTTGGCAAGCTTTGAACAACCATCAGCAG 422
QY 1094 AGGCGATACCTCAATGCTTTAGGACAGATCGCTCTGCTGCTGTGTGTCAGCGCAGGATTC 1153
Db |||||
423 ATGTTTAGCTACAGCTATGGACAGTGGCTTTTGGCAGCTGCCAAGTTGGAGGAGCT 482
QY 1154 CTCCTCGCTGCAGCAAGTTCTATAGGGTCACTCTGTAAACACAGCTTTACAAGACCTCAAAAT 1213
Db |||||
483 CCGCAGGAACAGCTGGCAGCTGCCAGATGAATGTAAACACAGCTTTACAAGACAGCGTTT 542
QY 1214 CTACAGGTTCT-----GATTAATAACACAGATATCAGCAGGTTATGATGCTTACAAAT 1267
Db |||||
543 CTTGCACTTCTTCAGCTCTTTATGTCAGCAGCAGCTTTCCGATGGATATTTGCTTTACAAA 602
QY 1268 CCATCAATGATGCTATGTTAGGACAGAAATGATGGGCTCGTGATGTATAAACAATG 1327
Db |||||
603 CACTGAATCTTTATATTCGAAGACAGAGCGCGTGCAGTCAGCTATTATGTCAAACTG 662
QY 1328 TAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCACGAACAGAGCTCGAGGACAG 1387
Db |||||
663 CAATCCCGCGCTTTCAGAGAGCGTTTCTCGTTCTGGCAGTAGAAAGTCAAGGACGCAGTG 722
QY 1388 AAAAAACAGATCAAGCCCTCGTAGGATGATTTCTGCAATAGCAGAACTCTTTGGAGATG 1447
Db |||||
723 CAGATGTCAGCAAGAGACGACGAAACTATTTGTCAGAGATAGCCAAACGTTAGTGATG 782
QY 1448 TCTATAGTCAGGTTTCGGCACTCAATCTGTAAATGACAGATCACTCAGTCGAATCCTCAAG 1507
Db |||||
783 TATATAGCGGTTACAGGTTCTTGATTTCTTATGATGCTACGATTTGTGAGCAATCCGCAAG 842

```

```

QY 1508 CGAATAATGAGGAGATCAGACAAAAGCTTACATCGGCAGTGACAAAAGCTCCACAGTTTG 1567
Db |||||
843 CAATCAAGAAAGAGATTATGCAAGAGCTCACGGCATCTATTAGCAAAAGCTCCACATTTG 902
QY 1568 GTATCCCTTATGTGCAACTTTCTTAATGACTCTWACACAGAAAGTTTCATAGTAAATTAGAAA 1627
Db |||||
903 GGTATCCCTGCTGTTTCAGAAATCTGTGGATAGCTTGCAGAAAGTTTGTGTCACAAATTGGAAA 962
QY 1628 GTTTGTTTCTGAAGATCTAGACACAGAGCTGAAATAAAAGCACTTTCTCTTTGAAAACGA 1687
Db |||||
963 GAGAGTTTCTTATGTTGGGAACTAGTCTCGCAGAAATCTCAAGAGAAATGGTTTAGAAAAC 1022
QY 1688 ACTCCTTGTATTATCAGCAGGTCGTGGTCAATATCGGCTCTCTATATTTCTGTTTATCTCC 1747
Db |||||
1023 AGCCCGCTTTCATTCAACAGAGTTTGGTAAACATTTCTCTATTTCTCTGTTTATCTTT 1082
QY 1748 AATAAC 1753
Db |||||
1083 CTTAAC 1088

```

Search completed: March 24, 2004, 01:57:11
Job time : 7202.51 secs

C;Accession: G71490
 R;Stephens, R.S.; Kalman, S.; Iammell, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
 A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
 A;Reference number: A71570; MUID:9900809; PMID:9784136
 A;Accession: G71490
 A;Molecule type: DNA
 A;Residues: 1-647 <ARN>
 A;Cross-references: GB:AE001333; GB:AE001273; NID:G3329068; PIDN:AAC6826.1; PID:G332906
 A;Experimental source: serotype D, strain UW-3/Cx
 C;Comment: This sequence was originally identified as homologous to part of a sequence (PIR:E72042).
 C;Genetics:
 A;Gene: CT622

Query Match 29.6%; Score 843; DB 2; Length 647;
 Best Local Similarity 35.5%; Pred. No. 2.3e-35;
 Matches 213; Conservative 113; Mismatches 243; Indels 31; Gaps 14;

QY 1 MSIAKLGIGASSNSSSTSSRS-ADVDSTTATPTPPPTFFDDYKQATYDTTFTSTSL 59
 DB 60 TSLRD--AAILKNSSGPTDLSQLEASTSTVTRVAARDYNEAKSNFDTAKSGLENATTL 117
 QY 60 ADIOAALVSLQDAVNIKDTAATDEE-TAIAAEWETKNADA VKVGAQITELAKYASDNOA 118
 DB 118 AEYETKMADLAALQDMERLAKQAEVTRKEALQEKQ----EVIDKLNQLVLEKQXQT 173
 QY 119 ILDSLGKUTSFDLQAALLQSVANNKAAELKEMQDNVPVPGKTPAIAQSLVDQTDATA 178
 DB 174 LKETLTITDSDAQIPAINSOLEINKNSADQIIKLEGGNI---SYEAVLTNAGEVIKASS 230
 QY 179 TQIEKGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAIATAKTQIAEAQK---F 235
 DB 231 EAGIKLQALQSLVDAGDSQAALVQAQONNSPDNIAATKLLIDAAETKVNELKOEHTGL 290
 QY 236 PDSPILOBAEOMVIAEQKDLNKPADGSDVNPQTTVGGSKQOGSSITGSI-----RV 288
 DB 291 TDSPLVKAKEGSIQAQKDIQIEKES-GSDPIVGFPS-GSAASAGSAGVALKSSNNSGRI 348
 QY 289 SMLDDAENETASILMSGFROMIHMNTENPDSQAQQLAAQAPAAK---AAGDSDAAA 345
 DB 349 SILLDDVDNEMAAIAQGFMSIEQFNVNPNATAKELQAMEAQLTAMSDQLVGADGELPA 408
 QY 346 ALADAQKALEALGKAGQOQGLNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKT--S 403
 DB 409 EQAIKDLAALQAL-KQPTDGLATAMGQVFAAAKVGGSAGTAGTVQMNVKQLYKTAFS 467
 QY 404 KSTGSDYKQTQISAGYDAYKSIDNAYGRANDATRDVINNVSTPALTTRSVPRARTEARG-P 462
 DB 468 STSSSSYAAALSDGYSAYKTLNLSYSESRG-VQSAISQTANPALSRVSRSIGIESQGRS 526
 QY 463 EKTQALARVLSGNRTLDGVYSQVSALQSVWQITQSNPQANNEIRQKLTSAVTKPQPF 522
 DB 527 ADASQAAETIVRDSQTLGDVYSRLQVLDLSMTITVSNPQVQNEEIMQKLTASIKAPQF 586
 QY 523 GYPVQLGNDSTQKFIKLESIFARGSRTAARIKALSFETNSLFTQQVLNIGSLYSYGL 582
 DB 587 GYPVQNSADSLQKFAQGLREFVDGERSLAESENARFKQPAFTQQVLNIVIASLFSYGL 646

RESULT 5
 T31110
 extracellular matrix binding protein - Abiotrophia defectiva (fragment)
 C;Species: Abiotrophia defectiva
 C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C;Accession: T31110
 R;Manganelli, R.; van de Rijn, I.
 Infect. Immun. 67, 50-56, 1999
 A;Title: Cloning and characterization of emb, a gene encoding the major adhesin of Streptococcus
 A;Reference number: Z20988; MUID:99081722; PMID:9864195
 A;Status: preliminary
 A;Molecule type: DNA

A;Residues: 1-2055 <MAN>
 A;Cross-references: EMBL:AF067776; NID:G3249002; PID:G3249003; PIDN:AAD03320.1
 C;Genetics:
 A;Gene: emb

Query Match 7.4%; Score 210; DB 2; Length 2055;
 Best Local Similarity 21.8%; Pred. No. 0.0098;
 Matches 129; Conservative 90; Mismatches 262; Indels 112; Gaps 24;

QY 10 ASSNSSSTSSRSADVDSTTATPTPPPTFFDDYK----- 43
 DB 1381 ATSNDDETAQNEGTQAINAVPQIPKAKT--DAKNAVTOADRKKDAIENDPNLTREEKV 1438
 QY 44 -----TQATAYDTTFTSTSLADIAQALVSLQDAVNIKID--TAATDEETAIAAEWET 94
 DB 1439 AAKAKVDAEAKKAKDAIDAAATSNADVTAKQNEGTKAINDVPTPTAKTDKNAVTOARADA 1498
 QY 95 KNADAVKVGACITELAKYASDNOAILDLSGLTSPDILLQALLOS--VANNKKAELIKE 152
 DB 1499 KK-DAIEKDPNLTREEKDA--KAKVDAEAKKAK-DAIDAATSNADVTAKQNEGTKAIND 1554
 QY 153 MQDNFVVPVPGKTPAIAQSLVDQTDATATQIEKGNAIIRDAYFA-----QGNASGAVEN 204
 DB 1555 VPQTP--AKTDA-KNAVTOAADAKKDAIEKDPNLTREEKDAKAKVDAEAKKAKDAIDA 1611
 QY 205 AKSN-----NSISNIDSAKAAATATAKTQIAEBAQK-KFPDPSILOBAQWVI 249
 DB 1612 ATSNADVTAKQAGKNAINAVPQTPPTAKTDKNAVTOAADAKKDAIENDANLTREEKDA 1671
 QY 250 QAEKDLNKPADGSDVNPQTTVGGSKQOGS-SIGST-RVSMELDDAEN--ETASILM 304
 DB 1672 KAKVDAEATKAKNAIDAAATSNADVTAKQNEGTKAINDVPTPTAKTDKNAVTOARATDKK 1731
 QY 305 SGFRQMIMFNTEN-----PDSQAQQLAAQALAAKAAAGDSDSAAALADAQKALEAAL 358
 DB 1732 SAI-----ENDPALTRREEKDAKAKVDAEATKAKNAIDAAATSNADVTAKQ----- 1776
 QY 359 GKASQOQGLNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKQTQISAGY 418
 DB 1777 -DAG--KNAINAVPQTTTAKTDKNAVQDAQDAKKAAIENDPALTRREEKDAKAKVDAEA 1833
 QY 419 DAYKSIDNAYG-----RARNDATERDVINNV-STPA-----LTRSVPRARTEARGPEKTDQA 468
 DB 1834 KKAKDAIDAAATSNADVTAKQAGKDAINAVPQTTTAKTDKNAVQDAQATDKSAIENDPA 1893
 QY 469 LARVISGNSRSLGVDYQVSALQSVWQITQSNPQANNEIRQKLTSAVTKPQP 521
 DB 1894 LTR--EEKDAVKAKVDAEAKKAKDAIDAAATSNADVTAKQTEG--TQAINAVPQ 1942

RESULT 6
 B89921
 hypothetical protein ebhA [imported] - Staphylococcus aureus (strain N315)
 C;Species: Staphylococcus aureus
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C;Accession: B89921
 R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, R.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A;Reference number: B89758; MUID:21311952; PMID:11418146
 A;Accession: B89921
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-6713 <KUR>
 A;Cross-references: GB:BA000018; PID:G33701232; PIDN:BAB42527.1; GSPDB:GN00149
 A;Experimental source: strain N315
 C;Genetics:
 A;Gene: ebhA

Query Match 7.2%; Score 206; DB 2; Length 6713;
 Best Local Similarity 21.5%; Pred. No. 0.074;

Matches	145;	Conservative	97;	Mismatches	279;	Indels	154;	Gaps	31;
---------	------	--------------	-----	------------	------	--------	------	------	-----

QY	5	DKLGIASSNSSSTSRSDYDS---	TTATAPTTPPPFTDDYKTKQAQAYDTITFS--TSL	59		
Db	2215	EKLAARAKQAKSDIGRLTDLNNAQRTAANAEDVQAPNL-----	AAVTAARAKKATSLNTAM	2269		
QY	60	ADIQAALVSLQDAVTNIK-----	DTAATDEETAIAAEWETKNADAVKVGQIIT	107		
Db	2270	GNLKHAAEKDNTKRSVNYTDAOPKQOAVDTAVTQAEITNA--	NGSNANETGVQALN	2327		
QY	108	ELAKYASDNOAILDSLGLKLSFDLQALQLQSVANNKKAARLLKEMQ--	DRNPVFEKTPAI	166		
Db	2328	QLNQAKND-----	LNGDNKVAQAKETAKRALASYNLNQAOSTAATSQIDNATTADVTA-	2382		
QY	167	AOSLVDTQDATATQIE--KQGNAI-----	RDAYF-AGQNASGAVENAKSN	209		
Db	2383	AQNTANELNTAMQOLQNGINDONTVKQOVNFTDADQCKDAYTNAVTAQOILDKANGON		2442		
QY	210	SIGNIDSAKAAIATAKTOIAEAAOKKFPDSPILQEAQWVIOAEKD-----	LKNIK	259		
Db	2443	-----	MTKAQVEAALNQVTTAKNAL-----	NGDANVRQAKSDAKANLGTLLHNNAQ	2489	
QY	260	PADGSDVNPFGTTVGG-----	SKQOG-----	SSIGS---	IRVSMLLDDAENETASI	302
Db	2490	KQDLTSQIEGATTGVNSVKTAKQADLDGAMQRLESAIANKDQTKASENYIDADFTKKTA		2549		
QY	303	LMSGFRQMIHWENTE--	NPDSQAACQELAAQARAAKAA--	GDDSAALADADAQKALEAAL	358	
Db	2550	FDNAITQAESYLKNDKHTNKDKQAVEAQISVSTENALNGDANLQCAKTEATQAINLT		2609		
QY	359	GKAGQQOQILNALQIASAAVWSAGVPPA--	AASSIGSSVKQLYKTSKSTGSDYKTOISA	416		
Db	2610	QLNTPQK--	TALKQVNAARQVSGVTDLKNATSLLNNAMDQL----	KQAIGDHTTVAG	2662	
QY	417	GY-----	DAYSINDAYGRARN--	DATROVINNV--	TPALTRSVPRARTEARG-----	461
Db	2663	GNTYNASPDQKQGYTDYNAKNIIVNGSPNVTNAADVTAATQORVNAETSLNGDNLAT		2722		
QY	462	-PEKTDQALRV-----	ISGNRTLGDVYSQVSALQSVMQITQS-----	NP	501	
Db	2723	AKQAKDALRQMTSLSDAQKSIITQI----	DSATQTVGVQSVKDNATNLNDAMNQLRNS	2778		
QY	502	QANNEIRKQLTSATVPQPGPYVOLSNDSOKF---	IAKLESLFAEGSRTAEIKAL	558		
Db	2779	IANKDEVKA-----	SQPYVDADTQKNAYTAVTSAENIINATSOPTLDPSAV	2826		
QY	559	SFETNSLFIQOVLVN	573			
Db	2827	TQAANQVNTNKTALN	2841			

RESULT	7
S33441	EF protein - Streptococcus suis
C/Species:	Streptococcus suis
C/Date:	06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C/Accession:	S33441
R/Smith, H.E.; Reek, F.H.; Vecht, U.; Gielkens, A.L.J.; Smits, M.A.	
A/Submitted to the EMBL Data Library, May 1993	
A/Description:	Repeats in an extracellular protein of wek-pathogenic strains are absent
A/Reference number:	S33441
A/Accession:	S33441
A/Status:	Preliminary
A/Molecule type:	DNA
A/Residues:	1-1822 <SM1>
A/Cross-references:	EMBL:X71880; NID:g298031; PIDN:CAA50714.1; PID:g298032

Query Match	7.2%;	Score	205;	DB	2;	Length	1822;
Best Local Similarity	23.7%;	Pred. No.	0.015;				
Matches	145;	Conservative	92;	Mismatches	229;	Indels	156;
QY	37	PTFDDYKTAQO-----	TAYDTITFTSLADIAALVSLQDA--VTNIKTPAA	81			

Db	1265	PNUSDEGQAEIKKLTDAVAKTTLATIEDNADXTQEAKEAQAALADLEKAKETQKIADKAA	1324
Qy	82	TDDETAIA--AEWETKNADAVKVGAGITELAKYASNDQAILDSLGKUTSFDLIQAALLQS	139
Db	1325	IDRLTILVKGDELEATKQDAKNKIADAAAKBAIASPNLTDAAEKKTFDVA--E-1380	
Qy	140	VANNKA-----AELLKEMQDNVPVPGKTPAIAQSLVDOTDATA--TQTEKGNARDA	191
Db	1381	VAKANDAI SAATSPADVQKE--EDAGVA-----AIAEDVLDAAKQDAKNKIAKDAAAKEA	1434
Qy	192	YFAGONASGA-----VENAKSNSISNIDSAAAIATATKTQITAEAAQKFPDPSPI	241
Db	1435	IGSNPNLTDAAKKTFETDVAEAVAKANDAIS-----AATSPADVQKE--EDAGVA	1482
Qy	242	QEAQOMVIOAEKDLKNIKPADGSDVNPFGTVGGSKOOGSSIGSIRVSMLLDDAENETAS	301
Db	1483	AIAEDVLDAKQDAKN-KIAKESD-----AAKSAIDANPNLTDAKESAK	1526
Qy	302	ILMSGFRQIMHMENTENPDSPAQAQELAA-----QARAAKAAGDDSAALADAK--	352
Db	1527	-----KAVDADAKAATDAIDASTSPVEAQAEDKGVGSIADQVLDAAKQDAK	1573
Qy	353	----ALPAALGKA-----CQOQIINALGOIASAA--VVSAGVPPAAASS-----I	392
Db	1574	NKIAKEVAAAKEAIDANPNLSDAEKASKAVADAKATTDAIDASTSPVEAQAEDKGV	1633
Qy	393	GSSVKOLYKTSKTSQSDYKTOISAGVDAYKSIDAYGR---ARNDAQDWIN---NVSTP	446
Db	1634	GSTRQDVILDAK--QDAKNKIAKESDAKSAIDANPNLTDAKESAKKAVDADAKAATD	1690
Qy	447	ALTRSVPRARTERGPB-KTDQALARVI-----SGNSRTLGDVYQSVALQSVMQITQS	499
Db	1691	AIDAST--SPVEAQAEDKGVGAIAKILDAAKQDAKNKIKAKEAESAVIDSNPLTDA	1748
Qy	500	NPOANNEEIRQKLTSAVT-----KPPQFGYPYVQLSNDSTQKF	537
Db	1749	AKGAASEIDKAVEBAIVLINGVRTYQLEKILPMAALIKPAAKVTTPVD--PNNLTEKE	1807
Qy	538	IAKLESLEFAGS	549
Db	1808	IARIKAFLENN	1819
RESULT 8			
B90835			
probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain R1			
C:Species: Escherichia coli			
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001			
C:Accession: B90835			
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G			
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.			
DNA Res. 8, 11-22, 2001			
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen			
A;Reference number: A99629; PMID:21156231; PMID:11258796			
A;Accession: B90835			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-971 <HAY>			
A;Cross-references: GB:BA000007; PIDN:BA035073.1; PID:gl3361114; GSPDB:GN00154			
A;Experimental source: strain O157:H7, substrain R1MD 0509952			
C;Genetics:			
A;Gene: ECa1650			
Query Match 7.0%; Score 199; DB 2; Length 971;			
Best Local Similarity 20.6%; Pred. No. 0.013;			
Matches 133; Conservative 79; Mismatches 227; Indels 206; Gaps 20;			
Qy	9	IASSNSSSTSRADVD-----STTATPTPPP-----TFDD	41
Db	38	VASENPDEAGRYSMDEVGYQYVILLVEGPPSHAGTITVYEDSQPTLNDFLGAMTDD	97
Qy	42	YKTAQATYATIF-----TSTSLADIQAALVSLQDAVTNIKDTAAATDEETAAEWETK	95

Job time : 20.9063 secs

```

RESULT 15
C82206
methyl-accepting chemotaxis protein Vcl1403 [imported] - Vibrio cholerae (strain N16961)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: C82206
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.I.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406:477-483, 2000
A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: C82206
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <HEI>
A;Cross-references: GB:AE004218; GB:AE003852; NID:g9655881; PIDN:AAF94560.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: Vcl1403
A;Map position: 1

Query Match          6.3%; Score 179.5; DB 2; Length 641;
Best Local Similarity 20.2%; Pred. No. 0.074;
Matches 125; Conservative 110; Mismatches 202; Indels 181; Gaps 28;

Qy      14 SSSSTSRSDVDSTTATPTPTPTFDYKTKQAOT-----AYDTFTSTSLADIQAALV 67
Db      9 TSASSSPSPNSHSTSAVAPKTEA---DDLREQARTLARKQQAERVASAT--AELLAGLE 63

Qy      68 SLQDAVNIKDTAATDEETAIAEWE---TKNADAV--KVGQAITEAKYASDNOAILDS 122
Db      64 ESSAMTQIRSSM---EQIAYGABESSSATKQSETAVSQMNFINEQATYATQSKILTK 120

Qy      123 LGLKTSFDLQALLOALQSVANNKAAELLKENODNPVPGKTPAQAQSLVDQDTATATCIE 182
Db      121 LEQ-----DIVDA-----NKINMVSNVQ-----TSERQNDSVKRMIE 155

Qy      183 KDGNAIRDAYFAGONASGAVENAKSNNSISNI---DSAKAAIATAKTQIAEAQKFPDS 238
Db      156 LSEQA-----AKINDAVQVHIADQTNLLNALNAIEAGRAGKHGKGF 198

Qy      239 PIQEAQOMVIQ-AEKDLKNTK-----ADGSDVPNPGTTVGGSK-----QQGSSIGS-- 285
Db      199 AVVADTVRTLAEKAEKNAANIESLIKIDTNGAEMVSNVSGVSSQKAESEVQKGHWVNKQL 258

Qy      286 --IRVSM--LDDAE-----NETASILMSGFRQIMHMENTENPDQAQQLAAQA 332
Db      259 TLIRNEVDLVKDEVLKKADEMSSAAVIAKGSSEV-----SQAQEQSAACE 308

Qy      333 RAAKAAGDSDAAALADAQAALKAALGKAGQOQOQILNALGQIASAAVVSAGVPPAAASSI 392
Db      309 ESLK-----SLDQOQIALDGAVTAAQSLDELTDDEL-RTSTDIVKSSSEVAAAEEL 358

Qy      393 GSSVKQLYKTSKTSQYKTOISAGYDAYKINDAYGARNDATRDVNNVST--PALTR 450
Db      359 SAGIEEINRSSNEI-MGALNQISSG-----AQHMAKSVETGITSLTQ 399

Qy      451 SVPRARTEARGPEKTDQALARVIG---NSRTLGDVYKSOVSALQSVMTQITOS----- 499
Db      400 IEQARLCKERAESLKLKACEKMLGIEENKTTVDEM-----ILAITESTKAATENL 450

Qy      500 NPQANNEIRQKLTSAV-----TKPPQFGYPYVQLSNDSTQKFI 538
Db      451 NEMANIERISQIDKIVDGISNVSIQTAMLVANGAVEAARAGEYKGFVVSTD----- 504

Qy      539 AKLESFPAEGSRTAARIK 556
Db      505 --IQNLANDAAENAEQIK 520
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:24:24 ; Search time 50.4852 Seconds
(without alignments)
3262.843 Million cell updates/sec

Title: US-10-608-559-4

Perfect score: 2850

Sequence: 1 MSLADKLGIASSNSSSTSR.....SLFIQQVLVNGISYGLQ 583

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2844	99.8	583	AAV71955	Aay71955 5'-trunca
2	2844	99.8	651	AAV71954	Aay71954 Chlamydia
3	2844	99.8	651	AAU38899	Aau38899 C. pneumo
4	2844	99.8	651	ABB94277	ABb94277 Chlamydia
5	2844	99.8	651	ABB90532	ABb90532 Chlamydia
6	2844	99.8	651	ADD42665	ADD42665 Chlamydia
7	1870	65.6	490	AAV71957	Aay71957 C. pneumo
8	1866	65.5	452	AAV71956	Aay71956 3'-trunca
9	1844	64.7	478	AAV35358	Aay35358 Chlamydia
10	995	34.9	212	AAV35357	Aay35357 Chlamydia
11	843	29.6	647	ABB94283	ABb94283 Chlamydia
12	843	29.6	647	ADD43866	ADD43866 Chlamydia
13	825.5	29.0	660	AAU38921	Aau38921 C. tracho
14	825.5	29.0	660	ADD42726	ADD42726 Chlamydia
15	665.5	23.4	361	AAU38921	Aau38921 C. tracho
16	665.5	23.4	361	AAU38921	Aau38921 C. tracho
17	665.5	23.4	361	AAU38921	Aau38921 C. tracho
18	599.5	21.0	361	AAU38921	Aau38921 C. tracho
19	227.5	8.0	350	AAV37571	Aay37571 Chlamydia
20	214.5	7.5	1178	ABJ26182	ABj26182 Aspergill
21	211.5	7.4	1463	AAE20110	AAe20110 lactobaci
22	211.5	7.4	1831	ABU43109	ABu43109 Protein e
23	211.5	7.4	1870	ABJ19019	ABj19019 Pathogen
24	211	7.4	603	ABU44080	ABu44080 Protein e
25	207	7.3	1822	AAE27745	AAe27745 Extracell

ALIGNMENTS

RESULT 1

AAV71955
ID AAY71955 standard; protein; 583 AA.

XX AC AAY71955;

XX DT 26-MAR-2001 (first entry)

XX DE 5'-truncated Chlamydia pneumoniae 76 kDa protein.

XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
XX upper respiratory tract disease; bronchitis; sinusitis;
XX acute respiratory disease; cough; sore throat; hoarseness; fever;
XX vaccine; immunisation; treatment; truncation mutant; mutein.

XX OS Chlamydia pneumoniae.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 497 /note= "Encoded by ACT"

XX WO200066739-A2.

XX 09-NOV-2000.

XX 03-MAY-2000; 2000WO-CA000511.

XX 03-MAY-1999; 99US-0132270P.

XX 30-JUN-1999; 99US-0141276P.

XX (AVET) AVENTIS PASTEUR LTD.

XX Murdin AD, Oomen RP, Wang J, Dunn P;

XX WPI; 2000-687542/67.

XX N-PSDB; AAD02064.

XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful for vaccinating against Chlamydia infections.

XX Claim 16b; Page 100-102; 112pp; English.

XX The present sequence is 5'-truncated Chlamydia pneumoniae 76 kDa protein.

XX C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory tract disease, bronchitis, sinusitis and acute respiratory disease such as cough, sore throat, hoarseness, fever; and abnormal chest sounds on

Handwritten signature

QY 361 AGQQGILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKQTQISAGYDA 420
 DB 429 AGQQGILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKQTQISAGYDA 488
 QY 421 YKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPEKTDQALARVISGNSRTL 480
 DB 489 YKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPEKTDQALARVISGNSRTL 548
 QY 481 GDVYSQVSALOSVMQITQSNPQANNEEIRQKLTSAVTKPPQGYVQVOLSNDSTQKFTAK 540
 DB 549 GDVYSQVSALOSVMQITQSNPQANNEEIRQKLTSAVTKPPQGYVQVOLSNDSTQKFTAK 608
 QY 541 LESLFAEGSRTAAETKALSFEFNSLFIQOVLVNGISLYSGYLQ 583
 DB 609 LESLFAEGSRTAAETKALSFEFNSLFIQOVLVNGISLYSGYLQ 651

RESULT 6

ID ADD42665
 XX ADD42665 standard; protein; 651 AA.

AC ADD42665;

DT 15-JAN-2004 (first entry)

XX Chlamydia pneumoniae antigen protein SEQ ID NO:78.

XX Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;
 KW antinfertility; cardiant; antiarteriosclerotic; ophthalmological;
 KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
 KW tubal obstruction; infertility; male infertility; ocular infection;
 KW blindness; acute respiratory tract infection; atherosclerosis;
 KW coronary heart disease.

XX Chlamydropbila pneumoniae.

XX WOZ003041560-A2.

XX 22-MAY-2003.

XX 05-NOV-2002; 2002WO-US035624.

XX 06-NOV-2001; 2001US-00012256.

XX 05-DEC-2001; 2001US-00007693.

XX 15-JUL-2002; 2002US-00197220.

XX (COR) CORIXA CORP.

XX Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;
 PI Probst P;

XX WPI: 2003-441771/41.

XX N-PSDB; ADD42825.

XX New DNA and proteins comprising a portion of a Chlamydia antigen, useful
 PT for diagnosing or treating Chlamydial infections, particularly as
 PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic
 PT inflammatory disease.

XX Disclosure; SEQ ID NO 78; 275pp; English.

XX The present invention describes compounds and methods for diagnosing and
 CC treating Chlamydial infection. Chlamydia polynucleotide and protein
 CC sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,
 CC antiarteriosclerotic and ophthalmological activities, and can be used in
 CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
 CC compositions or methods from the present invention can be used for the
 CC serodiagnosis or treatment of Chlamydial infections, particularly in
 CC humans. The polynucleotides, proteins or compositions are particularly
 CC useful for stimulating an immune response in a patient, or for
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC Specifically, the polynucleotides, proteins or compositions are useful as
 CC vaccines for treating or preventing Chlamydial infections including

CC pelvic inflammatory disease (which results in tubal obstruction and
 CC infertility in women), male infertility, ocular infection (which may
 CC cause blindness), acute respiratory tract infections, atherosclerosis, or
 CC coronary heart disease. The present sequence is used in the
 CC exemplification of the present invention.

SQ Sequence 651 AA;

Query Match 99.8%; Score 2844; DB 7; Length 651;

Best Local Similarity 99.8%; Pred. No. 1.4e-176;

Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSLADKLGIIASSNSSSTSRSDVSTTATPTPPPTFDDYKTKAQATYPTFTSTSLA 60

DB 69 MSLADKLGIIASSNSSSTSRSDVSTTATPTPPPTFDDYKTKAQATYPTFTSTSLA 128

QY 61 DIQAALVSLQDAVNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKYASDQAIL 120

DB 129 DIQAALVSLQDAVNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKYASDQAIL 188

QY 121 DSLGKLTSLFLLQALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVDQDATATQ 180

DB 189 DSLGKLTSLFLLQALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVDQDATATQ 248

QY 181 IEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAATATATAKTOIAEAKKPPDSPI 240

DB 249 IEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAATATATAKTOIAEAKKPPDSPI 308

QY 241 LQEAQWVIOAEKDLKNIKPADGSDVPNPVTGGSKQOQSSIGSIRVSMLLDDAENETA 300

DB 309 LQEAQWVIOAEKDLKNIKPADGSDVPNPVTGGSKQOQSSIGSIRVSMLLDDAENETA 368

QY 301 SILMSGFQMTMHPENTENPDSQAQOELAAQARAAGDDSAALADAKALEAALGK 360

DB 369 SILMSGFQMTMHPENTENPDSQAQOELAAQARAAGDDSAALADAKALEAALGK 428

QY 361 AGQQGILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKQTQISAGYDA 420

DB 429 AGQQGILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKQTQISAGYDA 488

QY 421 YKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPEKTDQALARVISGNSRTL 480

DB 489 YKSINDAYGRANDATRDVINNVSTPALTRSVPRARTEARGPEKTDQALARVISGNSRTL 548

QY 481 GDVYSQVSALOSVMQITQSNPQANNEEIRQKLTSAVTKPPQGYVQVOLSNDSTQKFTAK 540

DB 549 GDVYSQVSALOSVMQITQSNPQANNEEIRQKLTSAVTKPPQGYVQVOLSNDSTQKFTAK 608

QY 541 LESLFAEGSRTAAETKALSFEFNSLFIQOVLVNGISLYSGYLQ 583

DB 609 LESLFAEGSRTAAETKALSFEFNSLFIQOVLVNGISLYSGYLQ 651

RESULT 7

AAAY71957

ID AAAY71957 standard; protein; 490 AA.

XX AAAY71957;

XX 26-MAR-2001 (first entry)

XX C. pneumoniae 76 kDa protein truncation mutant fusion protein.

XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;

XX upper respiratory tract disease; bronchitis; sinusitis;

XX acute respiratory disease; cough; sore throat; hoarseness; fever;

XX vaccine; immunisation; treatment; fusion protein; truncation mutant;

XX Chlamydropbila pneumoniae.

OS Synthetic.

XX Location/Qualifiers

FH Key


```

QY 361 AGOQGGILNALGQIASAAVVSAGV 384
Db 429 AGOQGGILNALGQIASAAVVSAGV 452

RESULT 9
AAY35358
ID AAY35358 standard; protein; 478 AA.
XX
AC AAY35358;
XX
XX 17-OCT-2003 (revised)
DT 13-SEP-1999 (first entry)
XX
XX Chlamydia pneumoniae surface exposed polypeptide.
DE
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
KW neutralising epitope.
XX
XX Chlamydoghila pneumoniae.
OS
XX
XX WO9927105-A2.
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB001890.
XX
XX 21-NOV-1997; 97FR-00014673.
PR 04-NOV-1998; 98US-0107078P.
XX
XX (GEST ) GENSET.
XX
XX Griffais R;
XX
XX WPI; 1999-357842/30.
XX
XX Genome sequence of Chlamydia pneumoniae.
PS Page 1156; Disclosure; 1912pp; English.
XX
XX AAY34584-Y35879 represent the proteins encoded by all the open reading
frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
pneumoniae causes respiratory disease such as pneumonia and bronchitis
and is thought to be a contributing factor in heart disease, sarcoidosis,
sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
polypeptides encoded by the open reading frames of the C. pneumoniae
genome (see AAY34584-Y35879) can be used in immunogenic compositions as
vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
be used as immunogenic compositions, especially where the vector directs
the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
-OCT-2003 to standardise OS field)
XX
XX Sequence 478 AA;
Query Match 64.7%; Score 1844; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.1e-111; Indels 0; Gaps 0;
Matches 379; Conservative 0; Mismatches 0;

QY 1 MSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTAQAYDTFTSTSLA 60
Db 71 MSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTAQAYDTFTSTSLA 130

QY 61 DIQALVSLQDQAVTNIKOTAADEETAIAAEWETKNADAVKVGQITELAKYASDNQAIL 120
Db 131 DIQALVSLQDQAVTNIKOTAADEETAIAAEWETKNADAVKVGQITELAKYASDNQAIL 190

QY 121 DSLGKLTSDLLQALLOSVAANNKAAELLKEMQNPVVPVPGKTPAIQSLVDQDTATATQ 180
Db 191 DSLGKLTSDLLQALLOSVAANNKAAELLKEMQNPVVPVPGKTPAIQSLVDQDTATATQ 250

QY 181 FEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATAKTQIAEAKKFPDPSPI 240

```

```

Db 251 FEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATAKTQIAEAKKFPDPSPI 310
QY 241 LOEAQMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 300
Db 311 LOEAQMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 370
QY 301 SILMSGFRQMIHMFENTENPDSQAAQCELAQAARAKAAGDDSSAAAALADAQKALEAALGK 360
Db 371 SILMSGFRQMIHMFENTENPDSQAAQCELAQAARAKAAGDDSSAAAALADAQKALEAALGK 430
QY 361 AGOQGGILNALGQIASAAV 379
Db 431 AGOQGGILNALGQIASAAV 449

RESULT 10
AAY35357
ID AAY35357 standard; protein; 212 AA.
XX
XX AAY35357;
XX
XX 17-OCT-2003 (revised)
DT 13-SEP-1999 (first entry)
XX
XX Chlamydia pneumoniae transmembrane protein sequence.
DE
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
KW neutralising epitope.
XX
XX Chlamydoghila pneumoniae.
OS
XX
XX WO9927105-A2.
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB001890.
XX
XX 21-NOV-1997; 97FR-00014673.
PR 04-NOV-1998; 98US-0107078P.
XX
XX (GEST ) GENSET.
XX
XX Griffais R;
XX
XX WPI; 1999-357842/30.
XX
XX Genome sequence of Chlamydia pneumoniae.
PS Page 1155; Disclosure; 1912pp; English.
XX
XX AAY34584-Y35879 represent the proteins encoded by all the open reading
frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
pneumoniae causes respiratory disease such as pneumonia and bronchitis
and is thought to be a contributing factor in heart disease, sarcoidosis,
sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
polypeptides encoded by the open reading frames of the C. pneumoniae
genome (see AAY34584-Y35879) can be used in immunogenic compositions as
vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
be used as immunogenic compositions, especially where the vector directs
the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
-OCT-2003 to standardise OS field)
XX
XX Sequence 212 AA;
Query Match 34.9%; Score 995; DB 2; Length 212;
Best Local Similarity 99.0%; Pred. No. 6.5e-57;
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 380 VSAGVPPAAASSSTSSSVKQLYKPKSKTSGSDYKTKTQISAGYDAYKINDAYGRANDTRDV 439
Db 9 VSAGVPPAAASSSTSSSVKQLYKPKSKTSGSDYKTKTQISAGYDAYKINDAYGRANDTRDV 68

```



```

106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

```

```

CC serodagnosis or treatment of Chlamydial infections, particularly in
CC humans. The polynucleotides, proteins or compositions are particularly
CC useful for stimulating an immune response in a patient, or for
CC stimulating and/or expanding T cells specific for a Chlamydia protein.
CC Specifically, the polynucleotides, proteins or compositions are useful as
CC vaccines for treating or preventing Chlamydial infections including
CC pelvic inflammatory disease (which results in tubal obstruction and
CC infertility in women), male infertility, ocular infection (which may
CC cause blindness), acute respiratory tract infections, atherosclerosis, or
CC coronary heart disease. The present sequence is used in the
CC exemplification of the present invention.
XX
XX Sequence 660 AA;
XX
XX Query Match 29.0%; Score 825.5; DB 7; Length 660;
XX Best Local Similarity 34.8%; Pred. No. 3.2e-45;
XX Matches 212; Conservative 117; Mismatches 236; Indels 45; Gaps 15;
XX
XX 1 MSIALDKLGIASSNSSSTSRSS-ADVDSTTATATPTPPPTFDYKTOAQATYDTFTSTSL 59
XX 67 ISLRD--ALNKNSSPTDLSLQLEASTSTSTVTRVAADYDEAKSNFDTAKSGLENAKTL 124
XX 60 ADIQALVSLQDAVNIKDTATD-----EET-----AIAAEWETKNADAVKVGQITE 108
XX 125 AEYETKQADLMAALQDMERLANSDFSNHTEEVNINIKKALEAQKOTID-----KLNK 176
XX 109 LAKYASDNQAILDSLGKLTSPDLLQAALLQSVANNKKAELLKEMODNVPVPGKTPAIAQ 168
XX 177 LVTLQNKSLTEVLKTTDSADQIPAINSLQLEINKNASDQIITKDLERQNI---SYRAVIT 233
XX 169 SLVDQTDATATQTEKDGNAIRDYFAGQNAVENAKSNNSISNIDSAAIAATAKTQI 228
XX 234 NAGEVIKASSEAGIKLQALQSLVDAGDQSAVLQAQQNNSPDNIAATKELIDAAETKV 293
XX 229 ARAQK--FPSPILQEAQVVIQAEKOLKNIKPADGSDVPNPGTVGSKQOQSSIGS 285
XX 294 NELKQHTGLTDSPLVKAEEQISQAQKIOIKPS-GSDIPVGPS-GSAASAGSAGA 351
XX 286 I-----RVSMLLDDAENETASILMSGFROHMENTENPDSQAQQQLAAQARAAX-- 336
XX 352 LKSSNNSGRISLLDDVDNEMAIALQGFRLMEQFNVNPNPAKELQAMEAQLTAMSDQ 411
XX 337 -AAGDDSAALADAAQKALEAALGKAGQOQOQILNALGQIASAAVVSAGVPPAAASIGSS 395
XX 412 LVGADGELPAEIQAIKDALAQAAL-KQPSADGLATAMQVAFAAAKVGGSGAGTAGTVQWN 470
XX 396 VKQLYKT--SKSTGSDYKTOISAGYDAYKSIDNAYGRANDATRDVINNVSTPALTRSP 453
XX 471 VKQLYKTAFTSTSSSYAAALSDGYSAYKTLNLSYSESRSGVQSAISQTPANPALSRVS 529
XX 454 RARTEARG-PEKTDQALARVISGNSRTLDGVYSOVSALQSVMOITQSNPOANNEEIRQKL 512
XX 530 RSGIESQGRSADASQRAAETIVRDSQTLGDVYSRLQVLDLSLMTIIVSNPQANQEEIMQKL 589
XX 513 TSAVTKPQPGYPPVQLSNDSTQKFTAKLESFPAEGSRRTAAEIKALSFTNSLFIQOVLV 572
XX 590 TASISKAPQPGYPVQNSADSLQKFAAQLEREFVDGERSLAESQENAFRKPQAFIQOVLV 649
XX 573 NIGSLYSGYL 582
XX 650 NIASLFSGYL 659
XX
XX RESULT 15
XX AAB13695
XX ID AAB13695 standard; protein; 361 AA.
XX
XX AAB13695;
XX AC AC
XX XX
XX 02-FEB-2001 (first entry)
XX Chlamydia sp. protein # 6.
XX

```

KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;
KW coronary heart disease; antibacterial.
XX
XX Chlamydia sp.
XX
XX WO200034483-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 99WO-US029012.
XX
XX 08-DEC-1998; 98US-00208277.
XX 08-APR-1999; 99US-00288594.
XX 01-OCT-1999; 99US-00410568.
XX 22-OCT-1999; 99US-00426571.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
XX WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX comprises immunogenic portion of Chlamydia antigen, which comprises amino
XX acid sequence encoded by polynucleotide sequence.
XX
XX Disclosure; Page 251-253; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
XX proteins encoded by the nucleic acid sequences. The encoded proteins
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX proteins are useful for the serodiagnosis and treatment of Chlamydia
XX infection. Chlamydiae are intracellular bacterial pathogens that are
XX responsible for a wide variety of human infections. C. trachomatis
XX infection is one of the most common sexually transmitted diseases and
XX lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX and infertility. Trachoma due to ocular infection with C. trachomatis is
XX the leading cause of preventable blindness worldwide. C. pneumonia is a
XX major cause of acute respiratory tract infections in humans and is also
XX thought to play a role in the pathogenesis of atherosclerosis and
XX coronary heart disease. The present sequence is a protein isolated in the
XX present invention
XX
XX Sequence 361 AA;

Query Match 23.4%; Score 666.5; DB 3; Length 361;
Best Local Similarity 42.1%; Pred. No. 3.2e-35;
Matches 153; Conservative 62; Mismatches 131; Indels 17; Gaps 8;
QY 233 KFPDPSILOAEQVIOAEKDKLNKIPADGSDVDPNGTGVGGSKQQGSSIGSI----- 286
DB 2 QBIADSPLVKKAEEQINQAQQDIQTITPS-GLDIPVIGPS-GSAASAGSAGALKSSNNS 59
QY 287 -RVSMLLDDAENETASILMSGFROMHMENTENPDPSQAAQQLAAQARAAK---AAGDSDS 342
DB 60 GRISLLEDDVDNEMAIANQGRSMIEQFNVPNNPATAKELQAMEAQLTAMSDQLVGADGE 119
QY 343 AAAALADAKALEAALGKAGQQQGLNALGQIATAAVVSAGVPPAAASSIGSSVFQLYKT 402
DB 120 LPAEIQAIKDALAQAL-KQPSADGLATANGQVAFAAKVGGSGAGTACTVQNVKQLYKT 178
QY 403 --SKTSGDYKTQISAGVDYKAYSINDAYGRANDATRDVINNVSTPALTRSVPRARTEAR 460
DB 179 AFSSTSSSSYAAALSDGYSAKYLTLNLSXSESS-GVQSAISQTANPALSRVSRSRGIESQ 237
QY 461 G-PEKTDQALRVISGNRTLDGVYSOVSALOSVMQITCSNPQANNEIROKLTSAVTKP 519
DB 238 GRSADASQRAETIIVRDSQTLGVDVYSRQVLDLSMTIVSNPQANQEHIMQKLTASISKA 297
QY 520 PQFGYPYVQLSNDSTQKFIKLESLEFAEGSRTAABIKALSFPETNSLFTIQQVLVNIIGSLYS 579

Db 298 PQFGYPVQNSVDSLQKFAAQLEREFVDGERSLAESQENAFKQKQPAFTQQVIVNIASLFS 357
QY 580 GYL 582
Db 358 GYL 360
Search completed: March 23, 2004, 19:27:39
Job time : 52.4852 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 05:59:21 ; Search time 136.241 Seconds
(without alignments)
1108.119 Million cell updates/sec

Title: US-10-608-559-4
Perfect score: 2850
Sequence: 1 MSLADKLGIASSNSSSTSR.....SLFTQQVLWNTGSLYGYLQ 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2850	100.0	583	10	US-09-564-479-4
2	2844	99.8	651	9	US-09-841-132-405
3	2844	99.8	651	10	US-09-564-479-2
4	2844	99.8	651	10	US-09-841-260-78
5	2844	99.8	651	13	US-10-007-693-78
6	2844	99.8	651	15	US-10-312-273-13
7	1870	65.6	490	10	US-09-564-479-8
8	1866	65.5	452	10	US-09-564-479-6
9	1844	64.7	478	15	US-10-289-762-776
10	995	34.9	212	15	US-10-289-762-775
11	843	29.6	647	9	US-09-841-132-436
12	825.5	29.0	660	10	US-09-841-260-139
13	825.5	29.0	660	13	US-10-007-693-139
14	666.5	23.4	361	9	US-09-841-132-299
15	214.5	7.5	1178	14	US-10-128-714-8240

16	211.5	7.4	1463	9	US-09-971-536-69	Sequence 69, Appl
17	211.5	7.4	1831	12	US-10-282-122A-71033	Sequence 71033, A
18	211	7.4	603	12	US-10-282-122A-72004	Sequence 72004, A
19	203	7.1	5795	9	US-09-815-242-12610	Sequence 12610, A
20	202	7.1	6713	12	US-10-282-122A-43811	Sequence 43811, A
21	200.5	7.0	1123	12	US-10-282-122A-70581	Sequence 70581, A
22	199	7.0	2434	9	US-09-815-242-5835	Sequence 5835, Ap
23	199	7.0	6281	9	US-09-815-242-12996	Sequence 12996, A
24	197	6.9	2478	9	US-09-815-242-5816	Sequence 5816, Ap
25	197	6.9	2478	9	US-09-815-242-12967	Sequence 12967, A
26	196	6.9	560	9	US-09-815-242-13057	Sequence 13057, A
27	195.5	6.9	2271	12	US-10-282-122A-43924	Sequence 43924, A
28	194.5	6.8	2344	9	US-09-815-242-12713	Sequence 12713, A
29	194	6.8	6641	12	US-10-282-122A-70580	Sequence 70580, A
30	192	6.7	469	12	US-10-282-122A-70453	Sequence 70453, A
31	192	6.7	3533	12	US-10-282-122A-70177	Sequence 70177, A
32	190.5	6.7	2283	14	US-10-172-502-4	Sequence 4, Appli
33	187.5	6.6	2481	12	US-10-282-122A-43762	Sequence 43762, A
34	187.5	6.6	4776	12	US-10-282-122A-73678	Sequence 73678, A
35	187	6.6	2025	9	US-09-815-242-5703	Sequence 5703, Ap
36	187	6.6	2368	9	US-09-815-242-5635	Sequence 5635, Ap
37	187	6.6	2368	9	US-09-815-242-12389	Sequence 12389, A
38	187	6.6	3158	9	US-09-815-242-12611	Sequence 12611, A
39	185	6.5	3073	12	US-10-282-122A-49147	Sequence 49147, A
40	184	6.5	1142	15	US-10-369-493-10778	Sequence 10778, A
41	181	6.4	1215	9	US-09-815-242-5908	Sequence 5908, Ap
42	181	6.4	1269	9	US-09-815-242-13113	Sequence 13113, A
43	181	6.4	2086	9	US-09-815-242-5639	Sequence 5639, Ap
44	179.5	6.3	641	12	US-10-282-122A-77278	Sequence 77278, A
45	178.5	6.3	837	9	US-09-815-242-5883	Sequence 5883, Ap

ALIGNMENTS

RESULT 1

US-09-564-479-4
; Sequence 4, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-4

Query Match 100.0%; Score 2850; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 6e-205;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSLADKLGIASSNSSSTSR	ADVDSTTATATPTPPPTFDYKTAQATYDTFTSTSLA	60
Db	1	MSLADKLGIASSNSSSTSR	ADVDSTTATATPTPPPTFDYKTAQATYDTFTSTSLA	60
Qy	61	DIQAAVLSQDAVTN	IKDTAATDEETAIAEWETKNADAVKVGQITELAKYASDNQAIL	120
Db	61	DIQAAVLSQDAVTN	IKDTAATDEETAIAEWETKNADAVKVGQITELAKYASDNQAIL	120

Original

QY 121 DSLGKLTSLFDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQTDATQ 180
 DB 121 DSLGKLTSLFDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQTDATQ 180
 QY 181 IEKDNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 240
 DB 181 IEKDNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 240
 QY 241 LQAEQWVIOAEKDLKNIKPADGSDVNPCTTGGSKQOQSSIGSIRVSMMLDDAENETA 300
 DB 241 LQAEQWVIOAEKDLKNIKPADGSDVNPCTTGGSKQOQSSIGSIRVSMMLDDAENETA 300
 QY 301 SILMSGFQMLHMFNTENPDSCAAQOELAAQARAAGDSDSAAAALADAQKALEAALGK 360
 DB 301 SILMSGFQMLHMFNTENPDSCAAQOELAAQARAAGDSDSAAAALADAQKALEAALGK 360
 QY 361 AGQOQGILNALGQIASAAVVSAGVPPAAASIGSVKLYKTSKSTGSDYKTOISAGYDA 420
 DB 361 AGQOQGILNALGQIASAAVVSAGVPPAAASIGSVKLYKTSKSTGSDYKTOISAGYDA 420
 QY 421 YKSDINAYGRANDATRDVINNVSTPALTRSPVPRTEARGPEKTDQALARVSGNSRTL 480
 DB 421 YKSDINAYGRANDATRDVINNVSTPALTRSPVPRTEARGPEKTDQALARVSGNSRTL 480
 QY 481 GDVYSQVSALQSVNQITQSNPQANNEIRKLTSAVTKPPQFGYPYVQSLNSDSTQKFIK 540
 DB 481 GDVYSQVSALQSVNQITQSNPQANNEIRKLTSAVTKPPQFGYPYVQSLNSDSTQKFIK 540
 QY 541 LESLFAEGSRFAEIKALSFETNSLFIQOVLVNIIGSLYSGYLQ 583
 DB 541 LESLFAEGSRFAEIKALSFETNSLFIQOVLVNIIGSLYSGYLQ 583

RESULT 2
 US-09-841-132-405
 ; Sequence 405, Application US/09841132
 ; Patent No. US20020061848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.469C8
 ; CURRENT APPLICATION NUMBER: US/09/841,132
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 599
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 405
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-841-132-405

Query Match 99.8%; Score 2844; DB 9; Length 651;
 Best Local Similarity 99.8%; Pred. No. 2e-204;
 Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSLADKLGIASSNSSSTSRADVDSTTATPTPPPTFDDYKTOAQYATDTTSTSLA 60
 DB 69 MSLADKLGIASSNSSSTSRADVDSTTATPTPPPTFDDYKTOAQYATDTTSTSLA 128
 QY 61 DIQAALVSLQDVTNIKDTAATDEETAIAAEWETKNADAVKGAQITELAKYASDNQAIL 120
 DB 129 DIQAALVSLQDVTNIKDTAATDEETAIAAEWETKNADAVKGAQITELAKYASDNQAIL 188
 QY 121 DSLGKLTSLFDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQTDATQ 180
 DB 189 DSLGKLTSLFDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQTDATQ 248
 QY 181 IEKDNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 240
 DB 249 IEKDNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 308

QY 241 LQAEQWVIOAEKDLKNIKPADGSDVNPCTTGGSKQOQSSIGSIRVSMMLDDAENETA 300
 DB 309 LQAEQWVIOAEKDLKNIKPADGSDVNPCTTGGSKQOQSSIGSIRVSMMLDDAENETA 368
 QY 301 SILMSGFQMLHMFNTENPDSCAAQOELAAQARAAGDSDSAAAALADAQKALEAALGK 360
 DB 369 SILMSGFQMLHMFNTENPDSCAAQOELAAQARAAGDSDSAAAALADAQKALEAALGK 428
 QY 361 AGQOQGILNALGQIASAAVVSAGVPPAAASIGSVKLYKTSKSTGSDYKTOISAGYDA 420
 DB 429 AGQOQGILNALGQIASAAVVSAGVPPAAASIGSVKLYKTSKSTGSDYKTOISAGYDA 488
 QY 421 YKSDINAYGRANDATRDVINNVSTPALTRSPVPRTEARGPEKTDQALARVSGNSRTL 480
 DB 489 YKSDINAYGRANDATRDVINNVSTPALTRSPVPRTEARGPEKTDQALARVSGNSRTL 548
 QY 481 GDVYSQVSALQSVNQITQSNPQANNEIRKLTSAVTKPPQFGYPYVQSLNSDSTQKFIK 540
 DB 549 GDVYSQVSALQSVNQITQSNPQANNEIRKLTSAVTKPPQFGYPYVQSLNSDSTQKFIK 608
 QY 541 LESLFAEGSRFAEIKALSFETNSLFIQOVLVNIIGSLYSGYLQ 583
 DB 609 LESLFAEGSRFAEIKALSFETNSLFIQOVLVNIIGSLYSGYLQ 651

RESULT 3
 US-09-564-479-2
 ; Sequence 2, Application US/09564479
 ; Publication No. US20030095973A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: COHEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; CURRENT FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/132,270
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR APPLICATION NUMBER: 60/141,276
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-564-479-2

Query Match 99.8%; Score 2844; DB 10; Length 651;
 Best Local Similarity 99.8%; Pred. No. 2e-204;
 Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSLADKLGIASSNSSSTSRADVDSTTATPTPPPTFDDYKTOAQYATDTTSTSLA 60
 DB 69 MSLADKLGIASSNSSSTSRADVDSTTATPTPPPTFDDYKTOAQYATDTTSTSLA 128
 QY 61 DIQAALVSLQDVTNIKDTAATDEETAIAAEWETKNADAVKGAQITELAKYASDNQAIL 120
 DB 129 DIQAALVSLQDVTNIKDTAATDEETAIAAEWETKNADAVKGAQITELAKYASDNQAIL 188
 QY 121 DSLGKLTSLFDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQTDATQ 180
 DB 189 DSLGKLTSLFDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQTDATQ 248
 QY 181 IEKDNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 240
 DB 249 IEKDNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 308
 QY 241 LQAEQWVIOAEKDLKNIKPADGSDVNPCTTGGSKQOQSSIGSIRVSMMLDDAENETA 300

```

Db 309 LQEAQWVIAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSVKQYKTSKSGDYKTQISAGYDA 368
QY 301 SILMSGFRQMIHMFNTENPDSQAQQLAAQAAKAAAGDDSAALADAAKALEALGK 360
Db 369 SILMSGFRQMIHMFNTENPDSQAQQLAAQAAKAAAGDDSAALADAAKALEALGK 428
QY 361 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSGDYKTQISAGYDA 420
Db 429 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSGDYKTQISAGYDA 488
QY 421 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTL 480
Db 489 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTL 548
QY 481 GDVYSQVSALQVMQITQSNPOANNEEIRQKLTSAVTKPPQFGYPVQVLSNDSTQKFIK 540
Db 549 GDVYSQVSALQVMQITQSNPOANNEEIRQKLTSAVTKPPQFGYPVQVLSNDSTQKFIK 608
QY 541 LESLFAEGSRRTAAETKALSFTNSLFIQOVLVNIQSLYSGYLQ 583
Db 609 LESLFAEGSRRTAAETKALSFTNSLFIQOVLVNIQSLYSGYLQ 651

RESULT 4
US-09-841-260-78
; Sequence 78, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; FILE REFERENCE: 210121.515
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 78
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-78

Query Match 99.8%; Score 2844; DB 10; Length 651;
Best Local Similarity 99.8%; Pred. No. 2e-204;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSLADKLGIIASSNSSSTSRSDVSDTTATPTPPPTFDDYKTQAAQTAYDTFTSTSLA 60
Db 69 MSLADKLGIIASSNSSSTSRSDVSDTTATPTPPPTFDDYKTQAAQTAYDTFTSTSLA 128
QY 61 DIOAALVSLQDAVTNICKTAAATDEETAAEAETKNADAVKVGQITELAKYASNDQAIL 120
Db 129 DIOAALVSLQDAVTNICKTAAATDEETAAEAETKNADAVKVGQITELAKYASNDQAIL 188
QY 121 DSLGKLTSPDLLQALLOSIVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVDQTDATATQ 180
Db 189 DSLGKLTSPDLLQALLOSIVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVDQTDATATQ 248
QY 181 IEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATKTQIAEAQKPPDSP 240
Db 249 IEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATKTQIAEAQKPPDSP 308
QY 241 LQEAQWVIAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSVKQYKTSKSGDYKTQISAGYDA 300
Db 309 LQEAQWVIAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSVKQYKTSKSGDYKTQISAGYDA 368
QY 301 SILMSGFRQMIHMFNTENPDSQAQQLAAQAAKAAAGDDSAALADAAKALEALGK 360
Db 429 IEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATKTQIAEAQKPPDSP 428
QY 421 LQEAQWVIAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSVKQYKTSKSGDYKTQISAGYDA 480
Db 489 SILMSGFRQMIHMFNTENPDSQAQQLAAQAAKAAAGDDSAALADAAKALEALGK 548
QY 361 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSGDYKTQISAGYDA 608

```

```

Db 429 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSGDYKTQISAGYDA 488
QY 421 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTL 480
Db 489 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTL 548
QY 481 GDVYSQVSALQVMQITQSNPOANNEEIRQKLTSAVTKPPQFGYPVQVLSNDSTQKFIK 540
Db 549 GDVYSQVSALQVMQITQSNPOANNEEIRQKLTSAVTKPPQFGYPVQVLSNDSTQKFIK 608
QY 541 LESLFAEGSRRTAAETKALSFTNSLFIQOVLVNIQSLYSGYLQ 583
Db 609 LESLFAEGSRRTAAETKALSFTNSLFIQOVLVNIQSLYSGYLQ 651

RESULT 5
US-10-007-693-78
; Sequence 78, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; FILE REFERENCE: 210121.515C2
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 78
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-78

Query Match 99.8%; Score 2844; DB 13; Length 651;
Best Local Similarity 99.8%; Pred. No. 2e-204;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSLADKLGIIASSNSSSTSRSDVSDTTATPTPPPTFDDYKTQAAQTAYDTFTSTSLA 60
Db 69 MSLADKLGIIASSNSSSTSRSDVSDTTATPTPPPTFDDYKTQAAQTAYDTFTSTSLA 128
QY 61 DIOAALVSLQDAVTNICKTAAATDEETAAEAETKNADAVKVGQITELAKYASNDQAIL 120
Db 129 DIOAALVSLQDAVTNICKTAAATDEETAAEAETKNADAVKVGQITELAKYASNDQAIL 188
QY 121 DSLGKLTSPDLLQALLOSIVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVDQTDATATQ 180
Db 189 DSLGKLTSPDLLQALLOSIVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVDQTDATATQ 248
QY 181 IEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATKTQIAEAQKPPDSP 240
Db 249 IEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATKTQIAEAQKPPDSP 308
QY 241 LQEAQWVIAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSVKQYKTSKSGDYKTQISAGYDA 300
Db 309 LQEAQWVIAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSVKQYKTSKSGDYKTQISAGYDA 368
QY 301 SILMSGFRQMIHMFNTENPDSQAQQLAAQAAKAAAGDDSAALADAAKALEALGK 360
Db 369 SILMSGFRQMIHMFNTENPDSQAQQLAAQAAKAAAGDDSAALADAAKALEALGK 428
QY 361 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSGDYKTQISAGYDA 420
Db 429 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSGDYKTQISAGYDA 488
QY 421 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTL 480
Db 489 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTL 548
QY 481 GDVYSQVSALQVMQITQSNPOANNEEIRQKLTSAVTKPPQFGYPVQVLSNDSTQKFIK 540

```

Db 549 GDVYSQVSALQSVWQIIQSNPQANNEIRKLTSAVTKPQFGYPYVQLSNDSTQKFIK 608
Qy 541 LESLFAEGSRSTAABEIKALSFETNSLFIQQVLVNTGSLYSGYLQ 583
Db 609 LESLFAEGSRSTAABEIKALSFETNSLFIQQVLVNTGSLYSGYLQ 651

RESULT 6
US-10-312-273-13
; Sequence 13, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035W0
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 13
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-13

Query Match 99.8%; Score 2844; DB 15; Length 651;
Best Local Similarity 99.8%; Pred. No. 2e-204;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSLADKLGIASSSSSTSRGADVDSSTTATPTPPPTFDYKTAQATYDTTFTSLA 60
Db 69 MSLADKLGIASSSSSTSRGADVDSSTTATPTPPPTFDYKTAQATYDTTFTSLA 128
Qy 61 DIOAALVSLQDAVTKNTAATDEETAIAAEWETKNADAVKVGACITELAKYASDNQAIL 120
Db 129 DIOAALVSLQDAVTKNTAATDEETAIAAEWETKNADAVKVGACITELAKYASDNQAIL 188
Qy 121 DSLGKLTSPDLLQALLQSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVDQTDATQ 180
Db 189 DSLGKLTSPDLLQALLQSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVDQTDATQ 248
Qy 191 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFPDPSPI 240
Db 249 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFPDPSPI 308
Qy 241 LQEAQMWIQAEDKLNKIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 300
Db 309 LQEAQMWIQAEDKLNKIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 368
Qy 301 SILMSGFRQMIHMTENTENPDSQAQQLAAQARAARAAAGDDSAALADAKALEALGK 360
Db 369 SILMSGFRQMIHMTENTENPDSQAQQLAAQARAARAAAGDDSAALADAKALEALGK 428
Qy 361 AGCQGGILNALGQIASAAVWSAGVPPAAASSIGSSVKQLYKTSKSTGDKYKTQISAGYDA 420
Db 429 AGCQGGILNALGQIASAAVWSAGVPPAAASSIGSSVKQLYKTSKSTGDKYKTQISAGYDA 488

Qy 421 YKSINDAYGRANDATRDVINNVSTALTRESVPRATEARGPEKTKDQALARVISGNSRTL 480
Db 489 YKSINDAYGRANDATRDVINNVSTALTRESVPRATEARGPEKTKDQALARVISGNSRTL 548
Qy 481 GDVYSQVSALQSVWQIIQSNPQANNEIRKLTSAVTKPQFGYPYVQLSNDSTQKFIK 540
Db 549 GDVYSQVSALQSVWQIIQSNPQANNEIRKLTSAVTKPQFGYPYVQLSNDSTQKFIK 608

Qy 541 LESLFAEGSRSTAABEIKALSFETNSLFIQQVLVNTGSLYSGYLQ 583
Db 609 LESLFAEGSRSTAABEIKALSFETNSLFIQQVLVNTGSLYSGYLQ 651

RESULT 7
US-09-564-479-8
; Sequence 8, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COHEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: US2931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-8

Query Match 65.6%; Score 1870; DB 10; Length 490;
Best Local Similarity 99.7%; Pred. No. 1.3e-131;
Matches 385; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSLADKLGIASSSSSTSRGADVDSSTTATPTPPPTFDYKTAQATYDTTFTSLA 60
Db 69 MSLADKLGIASSSSSTSRGADVDSSTTATPTPPPTFDYKTAQATYDTTFTSLA 128
Qy 61 DIOAALVSLQDAVTKNTAATDEETAIAAEWETKNADAVKVGACITELAKYASDNQAIL 120
Db 129 DIOAALVSLQDAVTKNTAATDEETAIAAEWETKNADAVKVGACITELAKYASDNQAIL 188
Qy 121 DSLGKLTSPDLLQALLQSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVDQTDATQ 180
Db 189 DSLGKLTSPDLLQALLQSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVDQTDATQ 248
Qy 181 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFPDPSPI 240
Db 249 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFPDPSPI 308
Qy 241 LQEAQMWIQAEDKLNKIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 300
Db 309 LQEAQMWIQAEDKLNKIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 368
Qy 301 SILMSGFRQMIHMTENTENPDSQAQQLAAQARAARAAAGDDSAALADAKALEALGK 360
Db 369 SILMSGFRQMIHMTENTENPDSQAQQLAAQARAARAAAGDDSAALADAKALEALGK 428
Qy 361 AGCQGGILNALGQIASAAVWSAGVPP 386
Db 429 AGCQGGILNALGQIASAAVWSAGVLP 454

RESULT 8

US-09-564-479-6
; Sequence 6, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OOMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, FAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-6

Query Match
Best Local Similarity 65.5%; Score 1866; DB 10; Length 452;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLADKLGIIASSNSSSTSRSDVSTTATATPTPPPTFDDYKTOAQAYDTFTSTSLA 60
DB 69 MSLADKLGIIASSNSSSTSRSDVSTTATATPTPPPTFDDYKTOAQAYDTFTSTSLA 128
QY 61 DIQAAVLSQDAVTVNIKTAATDEETAAAEWETKNADAVKVGAGITELAKYASNQAIL 120
DB 129 DIQAAVLSQDAVTVNIKTAATDEETAAAEWETKNADAVKVGAGITELAKYASNQAIL 188
QY 121 DSLGKLTDFLLQALLQSVANNKAAELLKEMQNPVPGTTPAIAQSLVDQTDATATQ 180
DB 189 DSLGKLTDFLLQALLQSVANNKAAELLKEMQNPVPGTTPAIAQSLVDQTDATATQ 248
QY 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATQIAEAQKPPDSP 240
DB 249 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATQIAEAQKPPDSP 308
QY 241 LQEAQMVITQAEKDLKNIKPADGSDVPNPPTTGGSKQGGSSIGSRVSMMLDDAENETA 300
DB 309 LQEAQMVITQAEKDLKNIKPADGSDVPNPPTTGGSKQGGSSIGSRVSMMLDDAENETA 368
QY 301 SILMSGFRMIHMENTENPDSQAQELAAQARAARAAAGDDSAALADAQKALEALGK 360
DB 369 SILMSGFRMIHMENTENPDSQAQELAAQARAARAAAGDDSAALADAQKALEALGK 428
QY 361 AGQQGILNALGQIASAAVVSAGV 384
DB 429 AGQQGILNALGQIASAAVVSAGV 452

RESULT 9
US-10-289-762-776
; Sequence 776, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 776
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae

Query Match
Best Local Similarity 99.0%; Score 995; DB 15; Length 212;
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 380 VSAGVPPAAASSIGSVKQLYKTSKSTGSDYKTOISAGYDAYKSNIDAYGRANDTRDV 439
DB 9 VSAGVPPAAASSIGSVVNLKTSKSTGSDYKTOISAGYDAYKSNIDAYGRANDTRDV 68
QY 440 INNVSTPALTSVPRARTEARPEKTDQALARVIGSNRTLDGVYSQVSAQSVNQITQS 499
DB 69 INNVSTPALTSVPRARTEARPEKTDQALARVIGSNRTLDGVYSQVSAQSVNQITQS 128
QY 500 NPOANNEIROKLTSAVTKPPQFGYPYVQLSNDSTQKFIKLSFAGSGTAABIKALS 559
DB 129 NPOANNEIROKLTSAVTKPPQFGYPYVQLSNDSTQKFIKLSFAGSGTAABIKALS 188
QY 560 FETNSLFTQQVLMVIGSLYSGLQ 583

; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-776

Query Match
Best Local Similarity 64.7%; Score 1844; DB 15; Length 478;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLADKLGIIASSNSSSTSRSDVSTTATATPTPPPTFDDYKTOAQAYDTFTSTSLA 60
DB 71 MSLADKLGIIASSNSSSTSRSDVSTTATATPTPPPTFDDYKTOAQAYDTFTSTSLA 130
QY 61 DIQAAVLSQDAVTVNIKTAATDEETAAAEWETKNADAVKVGAGITELAKYASNQAIL 120
DB 131 DIQAAVLSQDAVTVNIKTAATDEETAAAEWETKNADAVKVGAGITELAKYASNQAIL 190
QY 121 DSLGKLTDFLLQALLQSVANNKAAELLKEMQNPVPGTTPAIAQSLVDQTDATATQ 180
DB 191 DSLGKLTDFLLQALLQSVANNKAAELLKEMQNPVPGTTPAIAQSLVDQTDATATQ 250
QY 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATQIAEAQKPPDSP 240
DB 251 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATQIAEAQKPPDSP 310
QY 241 LQEAQMVITQAEKDLKNIKPADGSDVPNPPTTGGSKQGGSSIGSRVSMMLDDAENETA 300
DB 311 LQEAQMVITQAEKDLKNIKPADGSDVPNPPTTGGSKQGGSSIGSRVSMMLDDAENETA 370
QY 301 SILMSGFRMIHMENTENPDSQAQELAAQARAARAAAGDDSAALADAQKALEALGK 360
DB 371 SILMSGFRMIHMENTENPDSQAQELAAQARAARAAAGDDSAALADAQKALEALGK 430
QY 361 AGQQGILNALGQIASAAV 379
DB 431 AGQQGILNALGQIASAAV 449

RESULT 10
US-10-289-762-775
; Sequence 775, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 775
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-775

Query Match
Best Local Similarity 99.0%; Score 995; DB 15; Length 212;
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 380 VSAGVPPAAASSIGSVKQLYKTSKSTGSDYKTOISAGYDAYKSNIDAYGRANDTRDV 439
DB 9 VSAGVPPAAASSIGSVVNLKTSKSTGSDYKTOISAGYDAYKSNIDAYGRANDTRDV 68
QY 440 INNVSTPALTSVPRARTEARPEKTDQALARVIGSNRTLDGVYSQVSAQSVNQITQS 499
DB 69 INNVSTPALTSVPRARTEARPEKTDQALARVIGSNRTLDGVYSQVSAQSVNQITQS 128
QY 500 NPOANNEIROKLTSAVTKPPQFGYPYVQLSNDSTQKFIKLSFAGSGTAABIKALS 559
DB 129 NPOANNEIROKLTSAVTKPPQFGYPYVQLSNDSTQKFIKLSFAGSGTAABIKALS 188
QY 560 FETNSLFTQQVLMVIGSLYSGLQ 583

ORGANISM: Aspergillus fumigatus
US-10-128-714-8240

Query Match 7.5%; Score 214.5; DB 14; Length 1178;
Best Local Similarity 23.5%; Pred. No. 5.2e-07;
Matches 148; Conservative 93; Mismatches 227; Indels 163; Gaps 29;

QY 16 SSTSRADYDSTTATAPTPPP-----PFDDYKTOAQAYDTFTSTSLADIQA-ALYS 68
DB 220 SSTSRSTNSTTVARPAKPMATATRP-----TTASKSLTAIKVSDDEAKLOS 271

QY 69 LQAVTNKIDTAATDEETAAAEWETKN-----ADAVKV--GAQ 105
DB 272 LQNKLESEATVA-----SLKALETVNEKLSQSVSQGASKENDEVASALRKEHAE 326

QY 106 ITTEL-AYASDNOAILDSLGL-TSPDLLQAAQLQSVANNKAA-----ELLK 151
DB 327 TAKLAARHAEQLQALQALDEAETKKELEKSIQLENAKAAAGDGDKTTAALEELK 386

QY 152 EMQDNFVPP-----GKTPAIAQSLVDOTATATQIEKGNAIRDA----- 191
DB 387 QSHQSOLETEKELABQKALAGVADQIAALQTLQESQKNGLEEAQKLEEKALALDSL 446

QY 192 --YFAGONASGAVENAKSNNSISNIDSAK-----AAATATAKQIABAQKFPDPSPILQEA 244
DB 447 DRELKGRDQ--VIENL--NMEMEKLSAKQEVEVRAAEETAKQSI SALEEKVAD-----L 496

QY 245 EQWVIOAEKDLKNIKPADGSDVNPGETTVGGSKQGGSSIGSIRVSMLLDDAENETASILM 304
DB 497 AAKLAQAE-----SARSQSSEETSLR-----LAEKKEVAEL-- 528

QY 305 SGFRQTHMENTENPDSQAQOELAAQARAAGDDSDAAALADADAKALEAALGK-AGQ 363
DB 529 ---KQAV-----AKSQEEL-----QAAREAAAKELSEKTAELAEAAHEAAVAKLAE 571

QY 364 QOGILNALGOIAGAAVVSAGVPPAAASSIGS-SVKQLYKTSKSTGSDYKTQISAGYDAYK 422
DB 572 HEG---ALASAAAHAQELSVAKQAESAGSMHSEELQKLREELDAAEAAAEEGREAIS 628

QY 423 SINDAYGRANDATRDVINNVSTPALTFRVPARTFARGPEKTDQALA-RVIGNSRTLG 481
DB 629 ELNAAH-QAELKALQDKLN-----ASEQALGEARQALEASVSSAQAVAVQVEIESLKEKVG 682

QY 482 DVYSQVSA---LQSVNQITQSNPQANNEELROKLTSAVTKPPQFGYPYVQLSNDSTQKF 537
DB 683 ALBAQLSTGQGELOSLIVEEVQSK-QAEABALHQNLTDFETQLKAKDAEKDEQLDLKEKA 741

QY 538 IA-----KLESLEFAGSGRTAAEIKALSPE 561
DB 742 AASEKALEEQLEAVAVAEHQAQALEAKAE 772

Search completed: March 24, 2004, 06:16:59
Job time : 138.241 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:39:25 ; Search time 10.7195 Seconds
(without alignments)
2831.945 Million cell updates/sec

Title: US-10-608-559-4

Perfect score: 2850

Sequence: 1 MSLADKLGIASSSSSTSR.....SLFTQQVLNIGSLYGVILQ 583

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	6.3	2541	1 TLN1_HUMAN	Q9V490 homo sapien
2	176	6.2	8545	1 ANCL1 CAEL	Q9N4M4 caenorhabdi
3	174	6.1	2541	1 TLN1_MOUSE	P26039 mus musculu
4	173.5	6.1	1140	1 Y96C YEAST	Q04893 saccharomyc
5	172	6.0	1306	1 MSB2 YEAST	P33334 saccharomyc
6	171.5	6.0	774	1 STE_LAMB	P03764 bacterioph
7	171	6.0	1120	1 STFR_ECOLI	P76072 escherichia
8	167	5.9	641	1 HRPK_PSESY	P41501 pseudomonas
9	165.5	5.8	2542	1 TLN2_HUMAN	Q9V4G6 homo sapien
10	163.5	5.7	881	1 PRV3 YEAST	P47033 saccharomyc
11	163.5	5.7	2090	1 N214_HUMAN	P35658 homo sapien
12	163	5.7	3644	1 MINT_MOUSE	Q62504 mus musculu
13	161.5	5.7	2492	1 TALAI_DICDI	P54633 dictyosteli
14	160	5.6	1161	1 DANA YEAST	P47179 saccharomyc
15	158	5.5	1562	1 SPAP_STRMU	P23504 streptococc
16	158	5.5	1565	1 PAC_STRMU	P11657 streptococc
17	158	5.5	1690	1 C196 DROME	Q9VJES drosophila
18	157.5	5.5	810	1 HTR4_HALSA	Q48317 halobacteri
19	156.5	5.5	810	1 HTR4_HALNI	Q9HP84 halobacteri
20	153	5.4	490	1 SE05_ECOLI	Q47500 escherichia
21	152.5	5.4	1528	1 SPAA_STRDO	P21979 streptococc
22	150.5	5.3	758	1 IMWT_HUMAN	Q16891 homo sapien
23	150	5.3	626	1 CEIB_ECOLI	P04479 escherichia
24	149.5	5.2	763	1 HTR2_HALNI	Q9HP81 halobacteri
25	149.5	5.2	764	1 HTR2_HALSA	P17410 halobacteri
26	149.5	5.2	1902	1 P2P_LACLC	P15293 lactococcus
27	149	5.2	516	1 P54_ENTFC	P13692 enterococcu
28	149	5.2	757	1 IMWT_MOUSE	Q8CAQ8 mus musculu
29	148.5	5.2	507	1 RRPD_CDOV	P06940 canine dist
30	148	5.2	1238	1 SBCC_RHOCA	O68032 rhodobacter
31	147.5	5.2	1011	1 SC4A_RICAF	Q9AJ83 rickettsia
32	147.5	5.2	1902	1 P1P_LACLC	P16271 lactococcus
33	146.5	5.1	962	1 IF2_NEIMA	Q9JTB5 neisseria m

RESULT 1

ID	TLN1_HUMAN	STANDARD;	PRT;	2541 AA.
AC	Q9V490; Q9NZQ2; Q9UHH8;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Talin 1.			
GN	TLN1 OR TLN.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1] SEQUENCE FROM N.A.			
RP	MAO L., Fan Y.H.;			
RA	"Complete cDNA sequence of human talin.";			
RT	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
RL	[2] SEQUENCE FROM N.A.			
RP	MEDLINE=20079169; PubMed=10610730;			
EX	Ben-Yosef T., Francomano C.A.;			
RA	"Characterization of the human talin (TLN) gene: genomic structure, chromosomal localization, and expression pattern.";			
RT	Genomics 62:316-319(1999).			
RL	CC -!- FUNCTION: Probably involved in connections of major cytoskeletal structures to the plasma membrane. High molecular weight cytoskeletal protein concentrated at regions of cell-substratum contact and, in lymphocytes, at cell-cell contacts (By similarity).			
CC	CC -!- SUBUNIT: Binds with high affinity to vinculin and with low affinity to integrins (By similarity).			
CC	CC -!- SIMILARITY: Contains 1 FERM domain.			
CC	CC -!- SIMILARITY: Contains 1 I/LWEQ domain.			
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	CC EMBL; AF078828; AAD13152.1; --			
CC	CC EMBL; AF177198; AAF23322.1; --			
CC	CC EMBL; AF178534; AAF27330.1; --			
CC	CC EMBL; AF178081; AAF27330.1; JOINED.			
CC	CC Genew; HGNC:11845; TLN1.			
CC	CC MIM; 186745; --			
CC	CC GO; GO:0005925; C:focal adhesion; NAS.			
CC	CC GO; GO:0005200; P:structural constituent of cytoskeleton; NAS.			
CC	CC GO; GO:0006928; P:cell motility; NAS.			
CC	CC GO; GO:0007016; P:cytoskeletal anchoring; NAS.			
CC	CC InterPro; IPR000299; Band 4.1.			
CC	CC InterPro; IPR002558; I/LWEQ.			
CC	CC Pfam; PF00373; Band_41; 1.			

ALIGNMENTS

34	146	5.1	625	1	VP35_BPAPS	Q9CLR3 bacterioph
35	146	5.1	1093	1	TMF1_HUMAN	P82094 homo sapien
36	145.5	5.1	707	1	Y99C YEAST	P47166 saccharomyc
37	145.5	5.1	1022	1	SCA4_RICCN	Q52658 rickettsia
38	145.5	5.1	1080	1	MRTB_MOUSE	P59759 mus musculu
39	145	5.1	3178	1	Y889_CAEEL	Q09624 caenorhabdi
40	145	5.1	3259	1	GOB1_HUMAN	Q14789 homo sapien
41	144.5	5.1	1509	1	MYSN_ACACA	P05659 acanthamoeb
42	144	5.1	1325	1	Y309_MYCPN	P75334 mycoplasma
43	144	5.1	1609	1	LMG1_HUMAN	P11047 homo sapien
44	143.5	5.0	991	1	SCA4_RICSI	Q9AJ77 rickettsia
45	143.5	5.0	5171	1	BPEA_HUMAN	O94833 homo sapien

```

DR Pfam; PF01608; I LWEO; 1.
DR ProDom; PD011820; ILWEO; 1.
DR SMART; SM00295; B41; 1.
DR SMART; SM00307; ILWEO; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS00057; FERM_3; 1.
DR PROSITE; PS00945; I LWEO; 1.
KW Structural protein; Cytoskeleton.
FT DOMAIN 86 403
FT DOMAIN 2340 2533
FT CONFLICT 824 824 (IN REF. 2; AAF27330).
FT CONFLICT 1227 1227 S -> L (IN REF. 1).
FT CONFLICT 1549 1549 P -> A (IN REF. 2; AAF27330).
FT CONFLICT 1604 1604 Q -> K (IN REF. 2; AAF27330).
FT CONFLICT 1701 1701 E -> Q (IN REF. 2; AAF27330).
FT CONFLICT 1718 1718 H -> N (IN REF. 2; AAF27330).
FT CONFLICT 1966 1966 A -> R (IN REF. 1).
FT CONFLICT 2256 2256 MISSING (IN REF. 2; AAF27330).
FT CONFLICT 2256 2256 MISSING (IN REF. 2; AAF27330).
SQ SEQUENCE 2541 AA; 269717 MW; 47FD9B22BCF47296 CRC64;

Query Match 6.3%; Score 180; DB 1; Length 2541;
Best Local Similarity 21.6%; Pred. No. 0.27;
Matches 148; Conservative 97; Mismatches 256; Indels 184; Gaps 32;

QY 49 AYDTFTST-----SLADI-----QAALV--SLQDVTNKTATDEE-----T 86
DB 801 ATDTILVTENIFSMGMDAGEMVQARILAQATSDLVNAIKADABGESDLENSRKLLSAA 860
QY 87 AIAEWETKNADAVK-----VGAQITELAKYASDNOAILDSLGLKTSFD 130
DB 861 KILADATAKMVEAAKAAHPDSEEQOORLEAREAEGLRMATNAAQNAIKKLVQRLEHA 920
QY 131 LLOAAA--LLOSVANNKAAELLKE-----MCDNVPVPGKTPAIAQSLVDQDQATATQ 180
DB 921 AKQAAASATQTIAAQAHAASPKASAGPOPLLQSCAVAFQIPLLQGV-----RGSQ 974
QY 181 IEKGNIRDAYFAGN-----ASGAVENAKSNNSINISAKAAIATATQI 228
DB 975 AQPDPSPAQLALIASQSFLQPGKMWAAAKASVPTIQDQASAMQLSCAKNLTGLAEAL 1034
QY 229 ABAQKPFDS--PILOAEQWVIO-AEKDLAKIIPA--DGSVDNPGTGVGSKQ--GSS 282
DB 1035 RTAAQKAQEAQCPLEMSALSVMQNLKDLQEVKAAARDGKLPBGTBTEKCTQDGLNS 1094
QY 283 IGS1--RVSMLLDAAE--NE-----TASILMSGFRQMIHMFN-----TENPDSQAQOEL 328
DB 1095 TKAVSSAIAQLLGEVAQGNENVAGTARDVAGGLRLSLAQAAAGVAALTSDDPAVQAIVLDT 1154
QY 329 AAQA-----RAAKAA--GDSAAAAADAAQKALEAALGKAGQ-----QQGLNALG 372
DB 1155 ASDVLDKASSLIEEAKAAGHFGDPESQORLAQAKAVTQALNRCVSCLPQGRDVDNALR 1214
QY 373 QI--ASAAVVSAGVPPA-----AASSIGSVKOLYKTSKST-----406
DB 1215 AVGDASKLLSDSLPPSTGTQEQASRLNEAAAGLNQATELVQASRTGTPQDLAPASGRF 1274
QY 407 GSDYKQTQISAGYDAYKSIDNAGRNDAIR--DVINN-----VSPALTRSVPRARTEAR 460
DB 1275 GQDFSTFLEAGV-----EMAGAPSQEDRAQVVSNLKGISMSSSLKLLAAKALSTDP 1327
QY 461 GPEKTDQ--ALARVIGNSRRTLGDVYSQ-----VSMQSVQMIOQTQSNPQANNE- 506
DB 1328 APNLSQAAAAARAVTDSINQITCTQQAQPGQKCDNALRELETVRELLENFPVQPINDM 1387
QY 507 -----ETROKLTSAVT-----KPPQFGYPVQVLSNDSTQKFAKLESFAE 547
DB 1388 SYFGCLDSVWNSKVLGEAWTGISQNAKNGNLPERG-----DAISTASKALCGTE 1438
QY 548 GSRTAEEIKALSFETNSLFIQQLV 572
DB 1439 AAAQAAYLVGVS-DPNSQAQQQLV 1462

```

RESULT 2

```

ID ANCL CABEL STANDARD; PRT: 8545 AA.
AC Q9N4M4; O61841; O61842;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Nuclear anchorage protein 1 (Anchorage 1 protein) (Nesprin homolog).
GN ANCL OR ZK973.6 OR T03A1.3/T03A1.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodermata; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE OF 1-1751 FROM N.A., FUNCTION, CHARACTERIZATION, AND
RP INTERACTION WITH F-ACTIN AND UNC-84.
RC STRAIN=Bristol N2;
RX MEDLINE=22364052; PubMed=12169658;
RA Starr D.A., Han M.;
RT "Role of ANCL-1 in tethering nuclei to the actin cytoskeleton.";
RL Science 298:406-409(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S., Goela D., Broy M.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 689-1243 FROM N.A.
RX MEDLINE=22296983; PubMed=12408964;
RA Zhang Q., Ragnauth C., Greener M.J., Shanahan C.M., Roberts R.G.;
RT "The nesprins are giant actin-binding proteins, orthologous to
RT Drosophila melanogaster muscle protein MSP-300.";
RL Genomics 80:473-481(2002).
CC -!- FUNCTION: Plays a central role in nuclear and mitochondrial
CC anchoring. Probably connects nuclei to the cytoskeleton by
CC interacting with unc-84 at the nuclear envelope and with F-actin
CC in the cytoplasm, creating a bridge across the nuclear envelope
CC between the cytoskeleton and the nucleus.
CC -!- SUBUNIT: Interacts with F-actin via its N-terminal domain.
CC Interacts indirectly with unc-84.
CC -!- SUBCELLULAR LOCATION: Type IV membrane protein (potential). The
CC largest part of the protein is cytoplasmic, while its C-terminal
CC part is associated either with the nuclear envelope, most probably
CC the outer nuclear membrane, or with mitochondrial membrane.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in all postembryonic
CC cells.
CC -!- DEVELOPMENTAL STAGE: First expressed in L1 larvae, and thereafter.
CC -!- DOMAIN: The large coiled coil domains are composed of 6 nearly
CC exact repeats of 903 residues. The last repeat is partial. These
CC repeats are conserved in Hawai (CB4856), Australia (AB4) and
CC Germany (RC301) strains. The length of the repeat may be
CC maintained because of a selective advantage to keep the protein
CC large and allow a single molecule to extend more than 0,5
CC micrometers.
CC -!- DOMAIN: The Klarsicht domain, which contains a potential
CC transmembrane domain, is essential for the nuclear envelope
CC targeting.
CC -!- SIMILARITY: Belongs to the nesprin family.
CC -!- SIMILARITY: Contains 1 actin-binding domain.
CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -!- SIMILARITY: Contains 1 Klarsicht domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/

```



```

RESULT 3
TLN1_MOUSE
ID TLN1_MOUSE STANDARD; PRT; 2541 AA.
AC P26039; Q8VEF0;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Talin 1.
GN TLN1 OR TLN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Fibroblast;
RX MEDLINE=91015390; PubMed=2120593;
RA Rees D.J.G., Ades S.A., Singer S.J., Hynes R.O.;
RT "Sequence and domain structure of talin.";
RL Nature 347:685-689(1990).
RN [2]
SEQUENCE OF 1603-2541 FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=2238625; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal
CC structures to the plasma membrane. High molecular weight
CC cytoskeletal protein concentrated at regions of cell-substratum
CC contact and, in lymphocytes, at cell-cell contacts.
CC -!- SUBUNIT: Binds with high affinity to vinculin and with low
CC affinity to integrins.
CC -!- PTM: Phosphorylated.
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -!- SIMILARITY: Contains 1 I/LWEQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X56123; CAA39588.1; -
CC DR EMBL; BC018557; AAH18557.1; -
CC PIR; S11661; S11661.
CC DR MGD; MGI:1099832; Tln.
CC GO; GO:0005925; C:focal adhesion; IDA.
CC InterPro; IPR000299; Band 4.1.
CC DR InterPro; IPR002558; ILWEQ.
CC DR Pfam; PF00373; Band 41; 1.
CC DR Pfam; PF01608; I_LWEQ; 1.
CC DR ProDom; PD011820; ILWEQ; 1.
CC SMART; SM00295; B41; 1.

```

```

SMART; SM00307; ILWEQ; 1.
DR PROSITE; PS00660; FERM 1; 1.
DR PROSITE; PS00661; FERM 2; 1.
DR PROSITE; PS00662; FERM 3; 1.
DR PROSITE; PS00663; FERM 4; 1.
DR PROSITE; PS00664; FERM 5; 1.
DR PROSITE; PS00665; FERM 6; 1.
DR PROSITE; PS00666; FERM 7; 1.
DR PROSITE; PS00667; FERM 8; 1.
DR PROSITE; PS00668; FERM 9; 1.
DR PROSITE; PS00669; FERM 10; 1.
DR PROSITE; PS00670; FERM 11; 1.
DR PROSITE; PS00671; FERM 12; 1.
DR PROSITE; PS00672; FERM 13; 1.
DR PROSITE; PS00673; FERM 14; 1.
DR PROSITE; PS00674; FERM 15; 1.
DR PROSITE; PS00675; FERM 16; 1.
DR PROSITE; PS00676; FERM 17; 1.
DR PROSITE; PS00677; FERM 18; 1.
DR PROSITE; PS00678; FERM 19; 1.
DR PROSITE; PS00679; FERM 20; 1.
DR PROSITE; PS00680; FERM 21; 1.
DR PROSITE; PS00681; FERM 22; 1.
DR PROSITE; PS00682; FERM 23; 1.
DR PROSITE; PS00683; FERM 24; 1.
DR PROSITE; PS00684; FERM 25; 1.
DR PROSITE; PS00685; FERM 26; 1.
DR PROSITE; PS00686; FERM 27; 1.
DR PROSITE; PS00687; FERM 28; 1.
DR PROSITE; PS00688; FERM 29; 1.
DR PROSITE; PS00689; FERM 30; 1.
DR PROSITE; PS00690; FERM 31; 1.
DR PROSITE; PS00691; FERM 32; 1.
DR PROSITE; PS00692; FERM 33; 1.
DR PROSITE; PS00693; FERM 34; 1.
DR PROSITE; PS00694; FERM 35; 1.
DR PROSITE; PS00695; FERM 36; 1.
DR PROSITE; PS00696; FERM 37; 1.
DR PROSITE; PS00697; FERM 38; 1.
DR PROSITE; PS00698; FERM 39; 1.
DR PROSITE; PS00699; FERM 40; 1.
DR PROSITE; PS00700; FERM 41; 1.
DR PROSITE; PS00701; FERM 42; 1.
DR PROSITE; PS00702; FERM 43; 1.
DR PROSITE; PS00703; FERM 44; 1.
DR PROSITE; PS00704; FERM 45; 1.
DR PROSITE; PS00705; FERM 46; 1.
DR PROSITE; PS00706; FERM 47; 1.
DR PROSITE; PS00707; FERM 48; 1.
DR PROSITE; PS00708; FERM 49; 1.
DR PROSITE; PS00709; FERM 50; 1.
DR PROSITE; PS00710; FERM 51; 1.
DR PROSITE; PS00711; FERM 52; 1.
DR PROSITE; PS00712; FERM 53; 1.
DR PROSITE; PS00713; FERM 54; 1.
DR PROSITE; PS00714; FERM 55; 1.
DR PROSITE; PS00715; FERM 56; 1.
DR PROSITE; PS00716; FERM 57; 1.
DR PROSITE; PS00717; FERM 58; 1.
DR PROSITE; PS00718; FERM 59; 1.
DR PROSITE; PS00719; FERM 60; 1.
DR PROSITE; PS00720; FERM 61; 1.
DR PROSITE; PS00721; FERM 62; 1.
DR PROSITE; PS00722; FERM 63; 1.
DR PROSITE; PS00723; FERM 64; 1.
DR PROSITE; PS00724; FERM 65; 1.
DR PROSITE; PS00725; FERM 66; 1.
DR PROSITE; PS00726; FERM 67; 1.
DR PROSITE; PS00727; FERM 68; 1.
DR PROSITE; PS00728; FERM 69; 1.
DR PROSITE; PS00729; FERM 70; 1.
DR PROSITE; PS00730; FERM 71; 1.
DR PROSITE; PS00731; FERM 72; 1.
DR PROSITE; PS00732; FERM 73; 1.
DR PROSITE; PS00733; FERM 74; 1.
DR PROSITE; PS00734; FERM 75; 1.
DR PROSITE; PS00735; FERM 76; 1.
DR PROSITE; PS00736; FERM 77; 1.
DR PROSITE; PS00737; FERM 78; 1.
DR PROSITE; PS00738; FERM 79; 1.
DR PROSITE; PS00739; FERM 80; 1.
DR PROSITE; PS00740; FERM 81; 1.
DR PROSITE; PS00741; FERM 82; 1.
DR PROSITE; PS00742; FERM 83; 1.
DR PROSITE; PS00743; FERM 84; 1.
DR PROSITE; PS00744; FERM 85; 1.
DR PROSITE; PS00745; FERM 86; 1.
DR PROSITE; PS00746; FERM 87; 1.
DR PROSITE; PS00747; FERM 88; 1.
DR PROSITE; PS00748; FERM 89; 1.
DR PROSITE; PS00749; FERM 90; 1.
DR PROSITE; PS00750; FERM 91; 1.
DR PROSITE; PS00751; FERM 92; 1.
DR PROSITE; PS00752; FERM 93; 1.
DR PROSITE; PS00753; FERM 94; 1.
DR PROSITE; PS00754; FERM 95; 1.
DR PROSITE; PS00755; FERM 96; 1.
DR PROSITE; PS00756; FERM 97; 1.
DR PROSITE; PS00757; FERM 98; 1.
DR PROSITE; PS00758; FERM 99; 1.
DR PROSITE; PS00759; FERM 100; 1.
DR PROSITE; PS00760; FERM 101; 1.
DR PROSITE; PS00761; FERM 102; 1.
DR PROSITE; PS00762; FERM 103; 1.
DR PROSITE; PS00763; FERM 104; 1.
DR PROSITE; PS00764; FERM 105; 1.
DR PROSITE; PS00765; FERM 106; 1.
DR PROSITE; PS00766; FERM 107; 1.
DR PROSITE; PS00767; FERM 108; 1.
DR PROSITE; PS00768; FERM 109; 1.
DR PROSITE; PS00769; FERM 110; 1.
DR PROSITE; PS00770; FERM 111; 1.
DR PROSITE; PS00771; FERM 112; 1.
DR PROSITE; PS00772; FERM 113; 1.
DR PROSITE; PS00773; FERM 114; 1.
DR PROSITE; PS00774; FERM 115; 1.
DR PROSITE; PS00775; FERM 116; 1.
DR PROSITE; PS00776; FERM 117; 1.
DR PROSITE; PS00777; FERM 118; 1.
DR PROSITE; PS00778; FERM 119; 1.
DR PROSITE; PS00779; FERM 120; 1.
DR PROSITE; PS00780; FERM 121; 1.
DR PROSITE; PS00781; FERM 122; 1.
DR PROSITE; PS00782; FERM 123; 1.
DR PROSITE; PS00783; FERM 124; 1.
DR PROSITE; PS00784; FERM 125; 1.
DR PROSITE; PS00785; FERM 126; 1.
DR PROSITE; PS00786; FERM 127; 1.
DR PROSITE; PS00787; FERM 128; 1.
DR PROSITE; PS00788; FERM 129; 1.
DR PROSITE; PS00789; FERM 130; 1.
DR PROSITE; PS00790; FERM 131; 1.
DR PROSITE; PS00791; FERM 132; 1.
DR PROSITE; PS00792; FERM 133; 1.
DR PROSITE; PS00793; FERM 134; 1.
DR PROSITE; PS00794; FERM 135; 1.
DR PROSITE; PS00795; FERM 136; 1.
DR PROSITE; PS00796; FERM 137; 1.
DR PROSITE; PS00797; FERM 138; 1.
DR PROSITE; PS00798; FERM 139; 1.
DR PROSITE; PS00799; FERM 140; 1.
DR PROSITE; PS00800; FERM 141; 1.
DR PROSITE; PS00801; FERM 142; 1.
DR PROSITE; PS00802; FERM 143; 1.
DR PROSITE; PS00803; FERM 144; 1.
DR PROSITE; PS00804; FERM 145; 1.
DR PROSITE; PS00805; FERM 146; 1.
DR PROSITE; PS00806; FERM 147; 1.
DR PROSITE; PS00807; FERM 148; 1.
DR PROSITE; PS00808; FERM 149; 1.
DR PROSITE; PS00809; FERM 150; 1.
DR PROSITE; PS00810; FERM 151; 1.
DR PROSITE; PS00811; FERM 152; 1.
DR PROSITE; PS00812; FERM 153; 1.
DR PROSITE; PS00813; FERM 154; 1.
DR PROSITE; PS00814; FERM 155; 1.
DR PROSITE; PS00815; FERM 156; 1.
DR PROSITE; PS00816; FERM 157; 1.
DR PROSITE; PS00817; FERM 158; 1.
DR PROSITE; PS00818; FERM 159; 1.
DR PROSITE; PS00819; FERM 160; 1.
DR PROSITE; PS00820; FERM 161; 1.
DR PROSITE; PS00821; FERM 162; 1.
DR PROSITE; PS00822; FERM 163; 1.
DR PROSITE; PS00823; FERM 164; 1.
DR PROSITE; PS00824; FERM 165; 1.
DR PROSITE; PS00825; FERM 166; 1.
DR PROSITE; PS00826; FERM 167; 1.
DR PROSITE; PS00827; FERM 168; 1.
DR PROSITE; PS00828; FERM 169; 1.
DR PROSITE; PS00829; FERM 170; 1.
DR PROSITE; PS00830; FERM 171; 1.
DR PROSITE; PS00831; FERM 172; 1.
DR PROSITE; PS00832; FERM 173; 1.
DR PROSITE; PS00833; FERM 174; 1.
DR PROSITE; PS00834; FERM 175; 1.
DR PROSITE; PS00835; FERM 176; 1.
DR PROSITE; PS00836; FERM 177; 1.
DR PROSITE; PS00837; FERM 178; 1.
DR PROSITE; PS00838; FERM 179; 1.
DR PROSITE; PS00839; FERM 180; 1.
DR PROSITE; PS00840; FERM 181; 1.
DR PROSITE; PS00841; FERM 182; 1.
DR PROSITE; PS00842; FERM 183; 1.
DR PROSITE; PS00843; FERM 184; 1.
DR PROSITE; PS00844; FERM 185; 1.
DR PROSITE; PS00845; FERM 186; 1.
DR PROSITE; PS00846; FERM 187; 1.
DR PROSITE; PS00847; FERM 188; 1.
DR PROSITE; PS00848; FERM 189; 1.
DR PROSITE; PS00849; FERM 190; 1.
DR PROSITE; PS00850; FERM 191; 1.
DR PROSITE; PS00851; FERM 192; 1.
DR PROSITE; PS00852; FERM 193; 1.
DR PROSITE; PS00853; FERM 194; 1.
DR PROSITE; PS00854; FERM 195; 1.
DR PROSITE; PS00855; FERM 196; 1.
DR PROSITE; PS00856; FERM 197; 1.
DR PROSITE; PS00857; FERM 198; 1.
DR PROSITE; PS00858; FERM 199; 1.
DR PROSITE; PS00859; FERM 200; 1.
DR PROSITE; PS00860; FERM 201; 1.
DR PROSITE; PS00861; FERM 202; 1.
DR PROSITE; PS00862; FERM 203; 1.
DR PROSITE; PS00863; FERM 204; 1.
DR PROSITE; PS00864; FERM 205; 1.
DR PROSITE; PS00865; FERM 206; 1.
DR PROSITE; PS00866; FERM 207; 1.
DR PROSITE; PS00867; FERM 208; 1.
DR PROSITE; PS00868; FERM 209; 1.
DR PROSITE; PS00869; FERM 210; 1.
DR PROSITE; PS00870; FERM 211; 1.
DR PROSITE; PS00871; FERM 212; 1.
DR PROSITE; PS00872; FERM 213; 1.
DR PROSITE; PS00873; FERM 214; 1.
DR PROSITE; PS00874; FERM 215; 1.
DR PROSITE; PS00875; FERM 216; 1.
DR PROSITE; PS00876; FERM 217; 1.
DR PROSITE; PS00877; FERM 218; 1.
DR PROSITE; PS00878; FERM 219; 1.
DR PROSITE; PS00879; FERM 220; 1.
DR PROSITE; PS00880; FERM 221; 1.
DR PROSITE; PS00881; FERM 222; 1.
DR PROSITE; PS00882; FERM 223; 1.
DR PROSITE; PS00883; FERM 224; 1.
DR PROSITE; PS00884; FERM 225; 1.
DR PROSITE; PS00885; FERM 226; 1.
DR PROSITE; PS00886; FERM 227; 1.
DR PROSITE; PS00887; FERM 228; 1.
DR PROSITE; PS00888; FERM 229; 1.
DR PROSITE; PS00889; FERM 230; 1.
DR PROSITE; PS00890; FERM 231; 1.
DR PROSITE; PS00891; FERM 232; 1.
DR PROSITE; PS00892; FERM 233; 1.
DR PROSITE; PS00893; FERM 234; 1.
DR PROSITE; PS00894; FERM 235; 1.
DR PROSITE; PS00895; FERM 236; 1.
DR PROSITE; PS00896; FERM 237; 1.
DR PROSITE; PS00897; FERM 238; 1.
DR PROSITE; PS00898; FERM 239; 1.
DR PROSITE; PS00899; FERM 240; 1.
DR PROSITE; PS00900; FERM 241; 1.
DR PROSITE; PS00901; FERM 242; 1.
DR PROSITE; PS00902; FERM 243; 1.
DR PROSITE; PS00903; FERM 244; 1.
DR PROSITE; PS00904; FERM 245; 1.
DR PROSITE; PS00905; FERM 246; 1.
DR PROSITE; PS00906; FERM 247; 1.
DR PROSITE; PS00907; FERM 248; 1.
DR PROSITE; PS00908; FERM 249; 1.
DR PROSITE; PS00909; FERM 250; 1.
DR PROSITE; PS00910; FERM 251; 1.
DR PROSITE; PS00911; FERM 252; 1.
DR PROSITE; PS00912; FERM 253; 1.
DR PROSITE; PS00913; FERM 254; 1.
DR PROSITE; PS00914; FERM 255; 1.
DR PROSITE; PS00915; FERM 256; 1.
DR PROSITE; PS00916; FERM 257; 1.
DR PROSITE; PS00917; FERM 258; 1.
DR PROSITE; PS00918; FERM 259; 1.
DR PROSITE; PS00919; FERM 260; 1.
DR PROSITE; PS00920; FERM 261; 1.
DR PROSITE; PS00921; FERM 262; 1.
DR PROSITE; PS00922; FERM 263; 1.
DR PROSITE; PS00923; FERM 264; 1.
DR PROSITE; PS00924; FERM 265; 1.
DR PROSITE; PS00925; FERM 266; 1.
DR PROSITE; PS00926; FERM 267; 1.
DR PROSITE; PS00927; FERM 268; 1.
DR PROSITE; PS00928; FERM 269; 1.
DR PROSITE; PS00929; FERM 270; 1.
DR PROSITE; PS00930; FERM 271; 1.
DR PROSITE; PS00931; FERM 272; 1.
DR PROSITE; PS00932; FERM 273; 1.
DR PROSITE; PS00933; FERM 274; 1.
DR PROSITE; PS00934; FERM 275; 1.
DR PROSITE; PS00935; FERM 276; 1.
DR PROSITE; PS00936; FERM 277; 1.
DR PROSITE; PS00937; FERM 278; 1.
DR PROSITE; PS00938; FERM 279; 1.
DR PROSITE; PS00939; FERM 280; 1.
DR PROSITE; PS00940; FERM 281; 1.
DR PROSITE; PS00941; FERM 282; 1.
DR PROSITE; PS00942; FERM 283; 1.
DR PROSITE; PS00943; FERM 284; 1.
DR PROSITE; PS00944; FERM 285; 1.
DR PROSITE; PS00945; FERM 286; 1.
DR PROSITE; PS00946; FERM 287; 1.
DR PROSITE; PS00947; FERM 288; 1.
DR PROSITE; PS00948; FERM 289; 1.
DR PROSITE; PS00949; FERM 290; 1.
DR PROSITE; PS00950; FERM 291; 1.
DR PROSITE; PS00951; FERM 292; 1.
DR PROSITE; PS00952; FERM 293; 1.
DR PROSITE; PS00953; FERM 294; 1.
DR PROSITE; PS00954; FERM 295; 1.
DR PROSITE; PS00955; FERM 296; 1.
DR PROSITE; PS00956; FERM 297; 1.
DR PROSITE; PS00957; FERM 298; 1.
DR PROSITE; PS00958; FERM 299; 1.
DR PROSITE; PS00959; FERM 300; 1.
DR PROSITE; PS00960; FERM 301; 1.
DR PROSITE; PS00961; FERM 302; 1.
DR PROSITE; PS00962; FERM 303; 1.
DR PROSITE; PS00963; FERM 304; 1.
DR PROSITE; PS00964; FERM 305; 1.
DR PROSITE; PS00965; FERM 306; 1.
DR PROSITE; PS00966; FERM 307; 1.
DR PROSITE; PS00967; FERM 308; 1.
DR PROSITE; PS00968; FERM 309; 1.
DR PROSITE; PS00969; FERM 310; 1.
DR PROSITE; PS00970; FERM 311; 1.
DR PROSITE; PS00971; FERM 312; 1.
DR PROSITE; PS00972; FERM 313; 1.
DR PROSITE; PS00973; FERM 314; 1.
DR PROSITE; PS00974; FERM 315; 1.
DR PROSITE; PS00975; FERM 316; 1.
DR PROSITE; PS00976; FERM 317; 1.
DR PROSITE; PS00977; FERM 318; 1.
DR PROSITE; PS00978; FERM 319; 1.
DR PROSITE; PS00979; FERM 320; 1.
DR PROSITE; PS00980; FERM 321; 1.
DR PROSITE; PS00981; FERM 322; 1.
DR PROSITE; PS00982; FERM 323; 1.
DR PROSITE; PS00983; FERM 324; 1.
DR PROSITE; PS00984; FERM 325; 1.
DR PROSITE; PS00985; FERM 326; 1.
DR PROSITE; PS00986; FERM 327; 1.
DR PROSITE; PS00987; FERM 328; 1.
DR PROSITE; PS00988; FERM 329; 1.
DR PROSITE; PS00989; FERM 330; 1.
DR PROSITE; PS00990; FERM 331; 1.
DR PROSITE; PS00991; FERM 332; 1.
DR PROSITE; PS00992; FERM 333; 1.
DR PROSITE; PS00993; FERM 334; 1.
DR PROSITE; PS00994; FERM 335; 1.
DR PROSITE; PS00995; FERM 336; 1.
DR PROSITE; PS00996; FERM 337; 1.
DR PROSITE; PS00997; FERM 338; 1.
DR PROSITE; PS00998; FERM 339; 1.
DR PROSITE; PS00999; FERM 340; 1.
DR PROSITE; PS01000; FERM 341; 1.
DR PROSITE; PS01001; FERM 342; 1.
DR PROSITE; PS01002; FERM 343; 1.
DR PROSITE; PS01003; FERM 344; 1.
DR PROSITE; PS01004; FERM 345; 1.
DR PROSITE; PS01005; FERM 346; 1.
DR PROSITE; PS01006; FERM 347; 1.
DR PROSITE; PS01007; FERM 348; 1.
DR PROSITE; PS01008; FERM 349; 1.
DR PROSITE; PS01009; FERM 350; 1.
DR PROSITE; PS01010; FERM 351; 1.
DR PROSITE; PS01011; FERM 352; 1.
DR PROSITE; PS01012; FERM 353; 1.
DR PROSITE; PS01013; FERM 354; 1.
DR PROSITE; PS01014; FERM 355; 1.
DR PROSITE; PS01015; FERM 356; 1.
DR PROSITE; PS01016; FERM 357; 1.
DR PROSITE; PS01017; FERM 358; 1.
DR PROSITE; PS01018; FERM 359; 1.
DR PROSITE; PS01019; FERM 360; 1.
DR PROSITE; PS01020; FERM 361; 1.
DR PROSITE; PS01021; FERM 362; 1.
DR PROSITE; PS01022; FERM 363; 1.
DR PROSITE; PS01023; FERM 364; 1.
DR PROSITE; PS01024; FERM 365; 1.
DR PROSITE; PS01025; FERM 366; 1.
DR PROSITE; PS01026; FERM 367; 1.
DR PROSITE; PS01027; FERM 368; 1.
DR PROSITE; PS01028; FERM 369; 1.
DR PROSITE; PS01029; FERM 370; 1.
DR PROSITE; PS01030; FERM 371; 1.
DR PROSITE; PS01031; FERM 372; 1.
DR PROSITE; PS01032; FERM 373; 1.
DR PROSITE; PS01033; FERM 374; 1.
DR PROSITE; PS01034; FERM 375; 1.
DR PROSITE; PS01035; FERM 376; 1.
DR PROSITE; PS01036; FERM 377; 1.
DR PROSITE; PS01037; FERM 378; 1.
DR PROSITE; PS01038; FERM 379; 1.
DR PROSITE; PS01039; FERM 380; 1.
DR PROSITE; PS01040; FERM 381; 1.
DR PROSITE; PS01041; FERM 382; 1.
DR PROSITE; PS01042; FERM 383; 1.
DR PROSITE; PS01043; FERM 384; 1.
DR PROSITE; PS01044; FERM 385; 1.
DR PROSITE; PS01045; FERM 386; 1.
DR PROSITE; PS01046; FERM 387; 1.
DR PROSITE; PS01047; FERM 388; 1.
DR PROSITE; PS01048; FERM 389; 1.
DR PROSITE; PS01049; FERM 390; 1.
DR PROSITE; PS01050; FERM 391; 1.
DR PROSITE; PS01051; FERM 392; 1.
DR PROSITE; PS01052; FERM 393; 1.
DR PROSITE; PS01053; FERM 394; 1.
DR PROSITE; PS01054; FERM 395; 1.
DR PROSITE; PS01055; FERM 396; 1.
DR PROSITE; PS01056; FERM 397; 1.
DR PROSITE; PS01057; FERM 398; 1.
DR PROSITE; PS01058; FERM 399; 1.
DR PROSITE; PS01059; FERM 400; 1.
DR PROSITE; PS01060; FERM 401; 1.
DR PROSITE; PS01061; FERM 402; 1.
DR PROSITE; PS01062; FERM 403; 1.
DR PROSITE; PS01063; FERM 404; 1.
DR PROSITE; PS01064; FERM 405; 1.
DR PROSITE; PS01065; FERM 406; 1.
DR PROSITE; PS01066; FERM 407; 1.
DR PROSITE; PS01067; FERM 408; 1.
DR PROSITE; PS01068; FERM 409; 1.
DR PROSITE; PS01069; FERM 410; 1.
DR PROSITE; PS01070; FERM 411; 1.
DR PROSITE; PS01071; FERM 412; 1.
DR PROSITE; PS01072; FERM 413; 1.
DR PROSITE; PS01073; FERM 414; 1.
DR PROSITE; PS01074; FERM 415; 1.
DR PROSITE; PS01075; FERM 416; 1.
DR PROSITE; PS01076; FERM 417; 1.
DR PROSITE; PS01077; FERM 418; 1.
DR PROSITE; PS01078; FERM 419; 1.
DR PROSITE; PS01079; FERM 420; 1.
DR PROSITE; PS01080; FERM 421; 1.
DR PROSITE; PS01081; FERM 422; 1.
DR PROSITE; PS01082; FERM 423; 1.
DR PROSITE; PS01083; FERM 424; 1.
DR PROSITE; PS01084; FERM 425; 1.
DR PROSITE; PS01085; FERM 426; 1.
DR PROSITE; PS01086; FERM 427; 1.
DR PROSITE; PS01087; FERM 428; 1.
DR PROSITE; PS01088; FERM 429; 1.
DR PROSITE; PS01089; FERM 430; 1.
DR PROSITE; PS01090; FERM 431; 1.
DR PROSITE; PS01091; FERM 432; 1.
DR PROSITE; PS01092; FERM 433; 1.
DR PROSITE; PS01093; FERM 434; 1.
DR PROSITE; PS01094; FERM 435; 1.
DR PROSITE; PS01095; FERM 436; 1.
DR PROSITE; PS01096; FERM 437; 1.
DR PROSITE; PS01097; FERM 438; 1.
DR PROSITE; PS01098; FERM 439; 1.
DR PROSITE; PS01099; FERM 440; 1.
DR PROSITE; PS01100; FERM 441; 1.
DR PROSITE; PS01101; FERM 442; 1.
DR PROSITE; PS01102; FERM 443; 1.
DR PROSITE; PS01103; FERM 444; 1.
DR PROSITE; PS01104; FERM 445; 1.
DR PROSITE; PS01105; FERM 446; 1.
DR PROSITE; PS01106; FERM 447; 1.
DR PROSITE; PS01107; FERM 448; 1.
DR PROSITE; PS01108; FERM 449; 1.
DR PROSITE; PS01109; FERM 450; 1.
DR PROSITE; PS01110; FERM 451; 1.
DR PROSITE; PS01111; FERM 452; 1.
DR PROSITE; PS01112; FERM 453; 1.
DR PROSITE; PS01113; FERM 454; 1.
DR PROSITE; PS01114; FERM 455; 1.
DR PROSITE; PS01115; FERM 456; 1.
DR PROSITE; PS01116; FERM 457; 1.
DR PROSITE; PS01117; FERM 458; 1.
DR PROSITE; PS01118; FERM 459; 1.
DR PROSITE; PS01119; FERM 460; 1.
DR PROSITE; PS01120; FERM 461; 1.
DR PROSITE; PS01121; FERM 462; 1.
DR PROSITE; PS01122; FERM 463; 1.
DR PROSITE; PS01123; FERM 464; 1.
DR PROSITE; PS01124; FERM 465; 1.
DR PROSITE; PS01125; FERM 466; 1.
DR PROSITE; PS01126; FERM 467; 1.
DR PROSITE; PS01127; FERM 468; 1.
DR PROSITE; PS01128; FERM 469; 1.
DR PROSITE; PS01129; FERM 470; 1.
DR PROSITE; PS01130; FERM 471; 1.
DR PROSITE; PS01131; FERM 472; 1.
DR PROSITE; PS01132; FERM 473; 1.
DR PROSITE; PS01133; FERM 474; 1.
DR PROSITE; PS01134; FERM 475; 1.
DR PROSITE; PS01135; FERM 476; 1.
DR PROSITE; PS01136; FERM 477; 1.
DR PROSITE; PS01137; FERM 478; 1.
DR PROSITE; PS01138; FERM 479; 1.
DR PROSITE; PS01139; FERM 480; 1.
DR PROSITE; PS01140; FERM 481; 1.
DR PROSITE; PS01141; FERM 482; 1.
DR PROSITE; PS01142; FERM 483; 1.
DR PROSITE; PS01143; FERM 484; 1.
DR PROSITE; PS01144; FERM 485; 1.
DR PROSITE; PS01145; FERM 486; 1.
DR PROSITE; PS01146; FERM 487; 1.
DR PROSITE; PS01147; FERM 488; 1.
DR PROSITE; PS01148; FERM 489; 1.
DR PROSITE; PS01149; FERM 490; 1.
DR PROSITE; PS01150; FERM 491; 1.
DR PROSITE; PS01151; FERM 492; 1.
DR PROSITE; PS01152; FERM 493; 1.
DR PROSITE; PS01153; FERM 494; 1.
DR PROSITE; PS01154; FERM 495; 1.
DR PROSITE; PS01155; FERM 496; 1.
DR PROSITE; PS01156; FERM 497; 1.
DR PROSITE; PS01157; FERM 498; 1.
DR PROSITE; PS01158;
```

```

GN YMR317W OR YMR924.09.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93(1997).
CC -!- DOMAIN: Contains many Ser/Thr-rich domain and repeats.
CC -!- SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; 254141; CA90835.1; -
DR GenOnline; 142998; -
DR SGD; S0004936; YMR317W.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;

Query Match 6.1%; Score 173.5; DB 1; Length 1140;
Best Local Similarity 16.9%; Pred. No. 0.2;
Matches 115; Conservative 127; Mismatches 277; Indels 163; Gaps 15;

QY 11 SSSSSSSSRADVDSTTATATPPP-----PTFDYKTAQTAVDTTFTS 58
DB 23 STSTTTPTTSTTSTTKVTTTPETIIVSSSTLVSVVPEFTSSSLSDTTASILS 82
QY 59 LADIQAAVLSQDVTNKDNTAATDEETAIAA-----EWETKNDAVKVG 103
DB 83 SESLVISFSLYSTSDISSTVNDVESSTGSPNSYALSSTNAQLSSITETDIS 142
QY 104 AQITELAYASNQAILDLGLKFTFDLLQALLOSANNKAAEL-----LKMQDN 156
DB 143 ALQTSPTQSSNG-----GGSSBPLGKSVLETTASSDPTTAVTSTFTLLDVSS 196
QY 157 PVVPGKTPAI-----AQSLVDQTDATATQIEKQNAIRDYFA-- 194
DB 197 PKISSGSAVTSVGTSTSDASKVFSSSTSDVSSLLSSTSSPASTISETLPSSITL 256
QY 195 -----GQNASGAVENAKNSNISNIDSAKAAIATAKTQIAEA-----QKFPDPSP 239
DB 257 SSPVSSEAPSATSSVSSEASSTSSVSSEAPLATSSVSSEAPSSSTSVVSSEAPS 316
QY 240 ILQEAQWVIAQEKDLKNIKPADGDVNPFGTGVGSKQGGSSIGSIRVSMLLDAENET 299
DB 317 SSSVSSEISTSSSTSSVSSEAPLATSSVSSEAPSSSTSSVSSEISTSSVSSEAPLAT 376
QY 300 ASILMSGFQMTHTMNTENPDQOAAQOELAAQAAKAGDDSAALADAKALEALG 359
DB 377 SSVVSS-----EAPSTSS--SVSSSEAPSTSSVSSEAPSSSTSSVSSEIST 423
QY 360 KAGQQQGIILNALQIASAASVAGVPPA-----AASSIGSS 395
DB 424 KS-----SVMSSEVSSTSSVSSEAPSAISSLASSRLFPSSKNTSVTSLVATEASVTS 479
QY 396 VKQLYKTSKTSQDYKTIQISAGDAYKINDAYGRANDATRDVNNVSTPALTSVPPA 455
DB 480 LRPSSETLAS-NSIIESSLSTGNTSVS-----TTTSAASSTLGGKSSVSGNRM 527
QY 456 RTEARGPEKTDQALARVISGNSRTL-----GDVYSQVSALQVMQI 496

```

```

DB 528 ATSKTSTSDLSKSVIFGNSVTTPSPASISLTASPLSPVSDITSSEASSISSMLA 587
QY 497 TQSNPQANNEIR-----OKLASAVTKPPQFGYPYVQLSNDSTQKFIKLE 542
DB 588 SSSAPSDNNSTIASASLIVTKNSVSVSIYSTTSSTTNSLNATSTSLSNKATAR 647
QY 543 SLFAGSGRTAAEIKALSFETNS 564
DB 648 SLTSTNATASAGNVPTGTFSSMS 669

RESULT 5
MSB2_YEAST STANDARD; PRT; 1306 AA.
AC P32334;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE MSB2 protein (Multicopy suppressor of bud emergence 2).
GN MSB2 OR YGR014W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92383951; PubMed=1514328;
RA Bender A., Pringle J.R.;
RT "A Ser/Thr-rich multicopy suppressor of a cdc24 bud emergence
RT defect."
RL Yeast 8:315-323 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieser M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII."
RL Yeast 13:1077-1090 (1997).
CC -!- PTM: O-glycosylated in the Ser/Thr-rich regions (Probable).
CC -!- SIMILARITY: SOME, TO YEAST HKR1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M77354; AAA34798.1; -
DR EMBL; Z72799; CRA96997.1; -
DR PIR; S25370; S25370.
DR GenOnline; 141326; -
DR SGD; S0003246; MSB2.
DR GO; GO:0005887; C:integral to plasma membrane; IMP.
DR GO; GO:0005034; F:osmensor activity; IMP.
DR GO; GO:0000283; P:establishment of cell polarity (sensu Sacch. .; IGI.
DR GO; GO:0006970; P:response to osmotic stress; IMP.
KW Transmembrane; Glycoprotein; Repeat.
FT DOMAIN 698 816 7 X 17 AA TANDEM REPEATS.
FT REPEAT 698 714 1.
FT REPEAT 715 731 2.
FT REPEAT 732 748 3.
FT REPEAT 749 765 4.
FT REPEAT 766 782 5.
FT REPEAT 783 799 6.
FT REPEAT 800 816 7.
SQ SEQUENCE 1306 AA; 133114 MW; 67D5D984D5CA4A6D CRC64;

Query Match 6.0%; Score 172; DB 1; Length 1306;
Best Local Similarity 21.4%; Pred. No. 0.29;

```


Query Match	5.7%;	Score 163.5;	DB 1;	Length 2090;
Best Local Similarity	21.1%;	Pred. No. 1.4;		
Matches 147;	Conservative 88;	Mismatches 262;	Indels 199;	Gaps 29;
QY	17	STSRADVSTSTATAATPP-----PTEFDDYKQAQTAYDTITFTSLADIQAALVSLQD	71	
DB	940	SPMKQQLRNFILAKRTTPVRSTAPASLSRASFSLQRYEYEDLDEVSTSSVSQSLD-ED	998	
QY	72	AVTNIKDTAATDEE-----		
DB	999	ARTSCKDDEAVVQAPRHPVWRTSPSQPSLLPHAAPFAKSHLVHGSSPGVMGTSA----	1054	
QY	94	TKNADAVKYGAQITELAKYASDNQALDSLGLKLTSPDLLQAAILLOSANNKKAELL---	150	
DB	1055	TSASKIIPQCADSTMLATKTVRHGAPSPHPIAPQOLAAALRRQNASQAPAVNTLUTS	1114	
QY	151	-----KEMQDNVPVPGKTPAIAGLSLVQDTDAT-----ATQTEKDGNAIRDA	191	
DB	1115	TLKNVPQVNVQELKKNPATP--STANGSSVPYSTAKTTPHPVLTPVAANQAKQSSLNSL	1172	
QY	192	YFAGQN-AGAVENAKSNNSISNIDSAKAIAIAKTCIAEAAQKFPDPSPILOBAEQMVIQ	250	
DB	1173	KPSGPTPASQOLSSGDKASGTAKIETAVTSTPASQGS---KPFSPSPSGTGFNGIIT	1229	
QY	251	AEKDLNKNIPADG-----SDVPNPGTIVGGSKQOGSSIGSRVSMLLDDAENETAS--	301	
DB	1230	PTPS-SNFTAAQCATSTKXESSQDPAFSSGGGSKPSYEALPESPSPSGITSANTTPGEP	1288	
QY	302	-----ILMSGFROMIHMFTNTPSQAAQ-----QELAAQA-----RAAKAAGDDS-	342	
DB	1289	AASSSRPVPASGTALSTTSKLETPPSKLGELLFPSSLAGETLGSFGLRVGQA--DDST	1346	
QY	343	-----AAALADAQAKLEA-----ALGKAGQOQGIILNALGQIATASAAVVSAGV	384	
DB	1347	KPTNKASSLSLSTQTKTIGSVPSGNGFTAPPVLGKHTEP-----PVTSGATTTSVA	1398	
QY	385	PAAASSIGSSV-KQLYKTSK-STG--SDYKTOISAGYDAYKSIINDAYGRARNDAT----	436	
DB	1399	PPAATSTSTAVFGLSPVTSAGSSGVIISFGTSSLSAGKTSF-----SFGSQQTNSTVPPS	1453	
QY	437	-----RDVINNVSPALTRSPVPRARTBARG-----PEKTOALARVI	473	
DB	1454	APPTTAATPLTPSPFTLSFGSLSSATTPLSPMSAGRSTEEATSSALPEK-----	1504	
QY	474	SGNSRTLGBDYVSQVSAQLSQVMQITQSNPOANNEIRKULTSAVTKPQFGYPYVQLSND	533	
DB	1505	PGDSEVSASNAALIEEQQA-QLPQAPQPTSD-----SVKKEPVLQAPAV--SNSG	1552	
QY	534	TQKFIAKLSLFPAG-----SRTAEIKALSF	560	
DB	1553	TRASSTSLVALSAEATPATTGVPDARTEAVPPASSF	1588	
RESULT 12				
ID	MINT_MOUSE	STANDARD;	PRT;	3644 AA.
AC	Q62504; Q80TN9; Q99PS4; Q9QZW2;			
DT	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Max2-interacting protein (SMART/HDAC1 associated repressor protein).			
OS	MINT OR SHARP OR KIAA0929.			
GN	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI	NCBI_TaxID=10090;			
RX				
RA	Newberry E.P., Latifi T., Towler D.A.;			
RP	SEQUENCE OF 49-3644 FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR			
RC	LOCATION, TISSUE SPECIFICITY, DNA-BINDING, AND INTERACTION WITH MSX2.			
RP	TISSUE=Testis;			
RA	MEDLINE=99379811; PubMed=10451362;			

RT "The RRM domain of MINT, a novel msx2 binding protein, recognizes and
 RT regulates the rat osteocalcin promoter.";
 RL Biochemistry 38:10678-10690(1999).
 RN [2]
 RP SEQUENCE OF 1-112 FROM N.A.
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
 RA Hashizume W., Hayashida K., Hirozane T., Hori F., Inotani K.,
 RA Itoh Y., Itoh M., Kagawa I., Kawai J., Kojima Y., Kondo S., Konno H.,
 RA Koya S., Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
 RA Ohno M., Ohsato N., Saito R., Sakazume N., Sano H., Sasaki D.,
 RA Sato K., Shibata K., Shiraki T., Tagami M., Takeda Y., Waki K.,
 RA Watahiki A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 69-3644 FROM N.A. (ISOFORM 2), AND VARIANTS THR-348;
 RP PHE-762; PHE-773 AND LEU-933.
 RC STRAIN=ICR; TISSUE=Brain;
 RA Sakamoto T., Gotou T., Isagawa Y., Mimura H., Kimura K., Kawaichi M.,
 RT "MINT/spen negatively regulates Notch signaling by inhibiting RBP-
 RT J/Su(H) activity.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 318-578 FROM N.A.
 RC TISSUE=Cochlea;
 RX MEDLINE=97237053; PubMed=9119401;
 RA Crozet F., El-Amraoui A., Blanchard S., Lenoir M., Ripoll C., Vago P.,
 RA Hamel C., Fizaes C., Levi-Acobas F., Depetris D., Mattei M.-G.,
 RA Weil D., Pujol R., Petit C.;
 RT "Cloning of the genes encoding two murine and human cochlear
 RT unconventional type I myosins.";
 RL Genomics 40:332-341(1997).
 RN [5]
 RP SEQUENCE OF 2598-3644 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Negase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 RN [6]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=22261914; PubMed=12374742;
 RA Oswald P., Kostecka U., Astrahantseff K., Bourteelle S., Dillinger K.,
 RA Zechner U., Ludwig L., Wilda M., Hameister H., Knoechel W., Liptay S.,
 RA Schmid R.M.;
 RT "SHARP is a novel component of the Notch/RBP-Jkappa signalling
 RT pathway.";
 RL EMBO J. 21:5417-5426(2002).
 RN [7]
 RP FUNCTION, AND TISSUE SPECIFICITY.
 RX MEDLINE=22483652; PubMed=12594956;
 RA Kuroda K., Han H., Tani S., Tanigaki K., Tun T., Furukawa T.,
 RA Taniguchi Y., Kurooka H., Hamada Y., Toyokuni S., Honjo T.;
 RT "Regulation of marginal zone B cell development by MINT, a suppressor
 RT of Notch/RBP-J signaling pathway.";
 RL Immunity 18:301-312(2003).
 CC -!- FUNCTION: Essential corepressor protein, which probably regulates
 CC different key pathways such as the Notch pathway. Negative
 CC regulator of the Notch pathway via its interaction with RBP5UH,
 CC which prevents the association between NOTCH1 and RBP5UH, and
 CC therefore suppresses the transactivation activity of Notch
 CC signaling. Blocks the differentiation of precursor B cells into
 CC marginal zone B cells. Probably represses transcription via the
 CC recruitment of large complexes containing histone deacetylase
 CC proteins. May bind both to DNA and RNA.
 CC -!- SUBUNIT: Interacts with NCOR2, HDAC1, RBBP4, MBD3 and
 CC TALL1. Interacts with the nuclear receptors RAR and PPARG.
 CC Interacts with RAR in absence of ligand. Bind to the steroid
 CC receptor RNA coactivator SRA (By similarity). Interacts with MSX2.
 CC Interacts with RBP5UH; this interaction may prevent the

CC interaction between RBP5UH and NOTCH1.
 CC -!- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q62504-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q62504-2; Sequence=VSP_008564;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Highly expressed in testis. Expressed at lower
 CC level in brain, lung, spleen, liver and kidney. Weakly expressed
 CC in cardiac and skeletal muscles and ovary. In spleen, it is
 CC expressed in follicular B-cells, while it is weakly expressed in
 CC marginal zone B-cells.
 CC -!- DOMAIN: The RID domain mediates the interaction with nuclear
 CC receptors.
 CC -!- DOMAIN: The SPOC domain, which mediates the interaction with
 CC NCOR2, is essential for the repressive activity (By similarity).
 CC -!- SIMILARITY: Belongs to the Spen family.
 CC -!- SIMILARITY: Contains 1 RID (receptor interacting) domain.
 CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
 CC -!- SIMILARITY: Contains 1 SPOC domain.
 CC -!- CAUTION: Ref.4 sequence differs from that shown due to multiple
 CC frame shifts and conflicts that create stop codons.
 CC -!- CAUTION: Ref.5 sequence differs from that shown due to what seems
 CC to be the presence of intronic sequence in the cDNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: BX726481; -; NOT_ANNOTATED_CDS.
 CC EMBL: AF156529; RAD55931.1; ALT_INIT.
 CC EMBL: AB055980; BAB32786.1; -.
 CC EMBL: Z78160; CAB01562.1; ALT_SEQ.
 CC EMBL: AK122402; BAC65684.2; ALT_SEQ.
 CC MGD: MG1:1391706; Mint.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC Pfam: PF00076; xtm; 3.
 CC SMART: SM00360; RRM; 3.
 CC PROSITE: PS0102; RRM; 4.
 CC PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 CC PROSITE: PS00917; SPOC_1; -.
 CC KW Transcription regulation; Repressor; Nuclear protein; DNA-binding;
 CC RNA-binding; Repeat; Coiled coil; Alternative splicing; Polymorphism.
 CC DOMAIN 1 574
 CC DNA-BINDING.
 CC FT DOMAIN 6 81
 CC RNA-BINDING (RRM) 1.
 CC FT DOMAIN 336 416
 CC RNA-BINDING (RRM) 2.
 CC FT DOMAIN 439 514
 CC RNA-BINDING (RRM) 3.
 CC FT DOMAIN 518 590
 CC RNA-BINDING (RRM) 4.
 CC FT DOMAIN 559 575
 CC COILED COIL (POTENTIAL).
 CC FT DOMAIN 822 850
 CC COILED COIL (POTENTIAL).
 CC FT DOMAIN 1185 1206
 CC COILED COIL (POTENTIAL).
 CC FT DOMAIN 1509 1544
 CC COILED COIL (POTENTIAL).
 CC FT DOMAIN 1607 1627
 CC COILED COIL (POTENTIAL).
 CC FT DOMAIN 2216 2704
 CC RID.
 CC FT DOMAIN 3478 3644
 CC SPOC.
 CC FT DOMAIN 2138 2462
 CC INTERACTION WITH MSX2.
 CC FT DOMAIN 2706 2845
 CC INTERACTION WITH RBP5UH.
 CC FT DOMAIN 125 277
 CC ARG-RICH.
 CC FT DOMAIN 236 326
 CC SER-RICH.
 CC FT DOMAIN 648 721
 CC TYR-RICH.
 CC FT DOMAIN 702 832
 CC ARG-RICH.
 CC FT DOMAIN 2101 2233
 CC ALA-RICH.
 CC FT DOMAIN 2377 2518
 CC PRO-RICH.
 CC FT DOMAIN 2950 3475
 CC PRO-RICH.
 CC FT VARSPLIC 618 640
 CC Missing (in isoform 2).
 CC /FTId=VSP_008564.
 CC I -> T.
 CC VARIANT 348 348

```
FT VARIANT 762 762 S -> F.
FT VARIANT 773 773 S -> F.
FT VARIANT 933 933 R -> L.
FT VARIANT 754 754 R -> G (IN REF. 3).
FT CONFLICT 1524 1524 D -> A (IN REF. 3).
FT CONFLICT 1560 1560 H -> Y (IN REF. 3).
FT CONFLICT 1570 1570 F -> L (IN REF. 3).
FT CONFLICT 1574 1574 Q -> G (IN REF. 3).
FT CONFLICT 1609 1609 Q -> R (IN REF. 3).
FT CONFLICT 1659 1659 I -> V (IN REF. 3).
FT CONFLICT 1669 1669 S -> F (IN REF. 3).
FT CONFLICT 1705 1705 V -> A (IN REF. 3).
FT CONFLICT 1815 1815 A -> V (IN REF. 3).
FT CONFLICT 2097 2097 G -> A (IN REF. 3).
FT CONFLICT 2201 2202 MISSING (IN REF. 3).
FT CONFLICT 2322 2322 A -> V (IN REF. 3).
FT CONFLICT 2385 2385 P -> Q (IN REF. 3).
FT CONFLICT 2502 2502 R -> K (IN REF. 3).
FT CONFLICT 2505 2505 E -> K (IN REF. 3).
FT CONFLICT 2519 2519 D -> N (IN REF. 3).
FT CONFLICT 2554 2554 T -> S (IN REF. 3).
FT CONFLICT 2679 2688 LVSTDPAGPVN -> VGEHPWARD (IN REF. 3).
FT CONFLICT 3010 3010 L -> P (IN REF. 3 AND 5).
FT CONFLICT 3644 3644 AA; 398750 MW; 9C7EC49A81A7DA4A CRC64;
SQ SEQUENCE

Query Match 5.7%; Score 163; DB 1; Length 3644;
Best Local Similarity 18.1%; Pred. No. 2.9;
Matches 115; Conservative 105; Mismatches 274; Indels 140; Gaps 16;

Qy 20 RSADVSTTATPTPPPTFDYKTOAQATVDTFTSTSLADIQALVSLQDVTNKT 79
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1620 KEAKQESFTHPTPEPAETKEPEK-----APVSAGLPAVTTVVTPPEA 1667

Qy 80 AATDETAIAEWETKNDAVKVGAQITELAKYASDNOAI-----LDSLGL 126
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1668 SSAPEKAEEAEAPSPAGEKPAEPVSEETKLSEPVSEVPEQSDVPDPPGDSRDSQ 1727

Qy 127 TSFDLLQALQSVANNKAAELKEMQDNVPVPGKTPALQSLVD----- 172
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1728 DS-----AALAPSAPOESAATDAVPCVNAEPLTGTTVTSQVSESSVDPKPSPLSKLTQ 1782

Qy 173 -QTDATATQIEK-----DGNALRDYFAQNASGAVENAKSNNSINIDSAKAATA 224
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1783 RSEBAEGEKVKPDTTPTTEPDATQNAVSEAQPPASEDVPAEKTRASKRS 1842

Qy 225 KTOIAEAQKFPDPSPILOEAE-----QMVIOAKDLKNIKPADG 263
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1843 KTSVQAAASVVEKPVTRKSERIDREKLKRSPPRGEAQKLELQWEAEKTRASKSSG 1902

Qy 264 SDVNPPTVGGSKQGSIGSIRVSMLLDDAENE-----TASILMSGFRQMIHM 313
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1903 GDTEHPEPSLPSRRNRNRSVATMTDHSRSPAKEFVEQPRVTRKRLRELQEA VVP 1962

Qy 314 FNT-----ENPDQAAQELAAQARAACA-----AGDSDAAALADAKALEAA 357
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1963 PTTPRGPPKTRRAEDGHEKREKPAETTPRPAEGWSPRSQSKAAAA-----GP 2013

Qy 358 LGKAGQOQGIINALQIAAAVWSAGVPPA-----AASSIGSVKLYKTSKGSDYKTQ 413
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2014 QGKGRNEQKVEAAEAQAQASTREGNPKSRGEREAASEPKRRDRDPSTDSKSPDTPFVE 2073

Qy 414 ISAGVDAYSNDAYGRARNDAITRVINNVSTPALTRSVPRPAREGPEKTDQALRVI 473
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2074 VLERKPPKTYKSKRGAR--STRSGMDRAHQRSLEMAARAAGQAADKE----- 2121

Qy 474 SGNRSLGDVYSOVALOSVMQITOSNPQ-ANN-----EIRQKLTSAVTKPPQFGYV 527
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2122 -----AGPAASPOESSPQKSGSSPOLANNPADPDRBAEESASASTAPPETQLAR 2175

Qy 528 QLSNDSTOKFTAKL-----ESLFAEGSRTAAEIKA 557
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2176 QIELEQAVQNIAKLPEPSAAAASKGTATATATAA 2209
```

```
RESULT 13
TALA_DICDI
ID TALA_DICDI STANDARD; PRT; 2492 AA.
AC P54633;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filopodin (Talin homolog).
GN TALA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Eukaryota; Dictyostelid; Dictyostelium.
OX NCBI_TaxID=44699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RX MEDLINE=95213284; PubMed=7698984;
RA Kreitmeier M., Gerisch G., Heizer C., Mueller-Taubenberger A.;
RT "A talin homologue of Dictyostelium rapidly assembles at the leading
RT edge of cells in response to chemoattractant.";
RL J. Cell Biol. 129:179-188(1995).
RN [2]
RP REVISIONS TO 2076; 2080; 2186 AND 2230.
RA Mueller-Taubenberger A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actin-binding protein that may be involved in the
CC control of cell motility and chemotaxis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; rapidly assembles at the
CC leading edge of cells in response to chemoattractant.
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -!- SIMILARITY: Contains 1 I/LWEQ domain.
CC PARTIAL, TO YEAST ROD1.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U14576; AAC4586.2; -.
CC DictyBase; DB0001399; talA.
CC InterPro; IPR00299; Band 4.1.
CC InterPro; IPR002558; ILWEQ.
CC Pfam; PF00373; Band 4.1; 1.
CC Pfam; PF01608; ILWEQ; 1.
CC ProDom; PD011820; ILWEQ; 1.
CC SMART; SM00295; B4.1; 1.
CC SMART; SM00307; ILWEQ; 1.
CC PROSITE; PS00660; FERM_1; 1.
CC PROSITE; PS00661; FERM_2; 1.
CC PROSITE; PS00057; FERM_3; 1.
CC PROSITE; PS00945; ILWEQ; 1.
CC Structural protein; Cytoskeleton; Actin-binding.
KW DOMAIN 84 365
FT DOMAIN 2298 2492 I/LWEQ.
SQ SEQUENCE 2492 AA; 268850 MW; 4B82F6BE6DE7E752 CRC64;

Query Match 5.7%; Score 161.5; DB 1; Length 2492;
Best Local Similarity 20.6%; Pred. No. 2.1;
Matches 121; Conservative 99; Mismatches 263; Indels 103; Gaps 20;

Qy 6 KLGIASSNSSSSRSADVSTTATPTPPPTFDYKTOAQATVDTFTSTSLADIQAA 65
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1982 KLLIVSSQLVGSRCQFHLGSATTKV-----ANTVSSLVKTAKIASVLAD 2029

Qy 66 LVSLQDVTNKTATDEBETAIAEWETKNDAVKVGAQITELAKYASDNOAILDSLGL 125
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2030 TTSSQDILSASKALSISQQMWLA---TKDAQRFKDA--TAFRSLGSAEVAEAVGQ 2083

Qy 126 -LTSFDLLQALQSVANNKAAELKEMQDNVPVPGKTPALQSLVDQTDAT----- 177
: : : : : : : : : : : : : : : : : : : : : : : : : :
```


[illegible]

Search completed: March 24, 2004, 05:56:59
Job time : 13.7195 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:47:55 ; Search time 43.9152 Seconds
(without alignments)
4188.690 Million cell updates/sec

Title: US-10-608-559-4
Perfect score: 2850
Sequence: 1 MSLAKLGIASSNSSSTR.....SLFIOQLVWIGSLGYLQ 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum Match 100%
Maximum DB seq length: 2000000000
Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2844	99.8	651	16 Q9Z7H7	Q9Z7H7 chlamydia p
2	1870	65.6	715	2 Q46166	Q46166 chlamydia p
3	1129	39.6	636	16 Q824X3	Q824X3 chlamydophi
4	843	29.6	647	16 Q84627	Q84627 chlamydia t
5	224.5	7.9	2310	16 Q8CMU7	Q8CMU7 staphylococ
6	210	7.4	2055	2 Q85472	Q85472 abiotrophia
7	206	7.2	6713	16 Q99U54	Q99U54 staphylococ
8	205	7.2	1822	2 Q07290	Q07290 streptococ
9	205	7.2	2016	5 Q9BIT0	Q9BIT0 plectreurus
10	202	7.1	6713	16 Q931R6	Q931R6 staphylococ
11	200.5	7.0	9439	16 Q8CP76	Q8CP76 staphylococ
12	200	7.0	1313	5 Q9NE28	Q9NE28 leishmania
13	200	7.0	9904	16 Q8NWQ6	Q8NWQ6 staphylococ
14	199.5	7.0	2275	16 Q8NUJ3	Q8NUJ3 staphylococ
15	199	7.0	973	16 Q8XDQ4	Q8XDQ4 escherichia
16	198	6.9	762	16 Q99IU6	Q99IU6 bradyrhizob

17	197	6.9	2478	2 Q9RL69	Q9RL69 staphylococ
18	197	6.9	2478	2 Q9LCH2	Q9LCH2 staphylococ
19	195.5	6.8	2271	16 Q99QY4	Q99QY4 staphylococ
20	194	6.8	2283	2 Q8VQ99	Q8VQ99 staphylococ
21	192	6.7	2167	2 Q845L8	Q845L8 mycoplasma
22	191.5	6.7	806	16 Q88Y67	Q88Y67 lactobacill
23	190.5	6.7	1545	16 Q9RDQ1	Q9RDQ1 streptomyce
24	190	6.7	856	5 Q86RM6	Q86RM6 plasmodium
25	189	6.6	855	5 Q86DH2	Q86DH2 plasmodium
26	189	6.6	1063	16 Q9CH86	Q9CH86 lactococcus
27	189	6.6	1365	2 Q49525	Q49525 mycoplasma
28	188	6.6	803	5 Q86RN4	Q86RN4 plasmodium
29	187.5	6.6	1795	16 Q9LCU9	Q9LCU9 staphylococ
30	187.5	6.6	2481	16 Q99QR6	Q99QR6 staphylococ
31	187.5	6.6	4776	16 Q97P71	Q97P71 streptococ
32	186	6.5	3360	16 Q88XB6	Q88XB6 lactobacill
33	185.5	6.5	1336	16 Q9ACV2	Q9ACV2 streptomyce
34	184.5	6.5	642	2 Q7X4Y6	Q7X4Y6 rhodospiril
35	184.5	6.5	855	5 Q86RM1	Q86RM1 plasmodium
36	183	6.4	625	5 Q86RN6	Q86RN6 plasmodium
37	182.5	6.4	3072	2 Q939N5	Q939N5 streptococ
38	182	6.4	842	5 Q86RN1	Q86RN1 plasmodium
39	182	6.4	3381	2 Q9KX33	Q9KX33 streptococ
40	181.5	6.4	2541	13 Q8AW10	Q8AW10 gallus gall
41	180.5	6.3	1327	2 Q9X7M2	Q9X7M2 staphylococ
42	180	6.3	2178	2 Q9KWR3	Q9KWR3 streptococ
43	180	6.3	2541	4 Q86YD0	Q86YD0 homo sapien
44	180	6.3	2550	4 Q9UPX3	Q9UPX3 homo sapien
45	179.5	6.3	641	16 Q9KS57	Q9KS57 vibrio chol

ALIGNMENTS

RESULT 1

ID	Q9Z7H7	PRELIMINARY;	PRT;	651 AA.
AC	Q9Z7H7			
DT	01-MAY-1999 (TREMELrel. 10, Created)			
DT	01-MAY-1999 (TREMELrel. 10, Last sequence update)			
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)			
DE	CHLPN 76 kDa HOMOLOG 1 (CT622) (Hypothetical protein)			
GN	CPN0728 OR CPJ0728 OR CP0018 OR CPB0756.			
OS	Chlamydia pneumoniae (Chlamydophila pneumoniae).			
OC	Bacteria; Chlamydiae; Chlamydophila pneumoniae; Chlamydiales; Chlamydiaceae; Chlamydophila.			
OX	NCBI_TaxID=83358;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CNL029;			
RX	MEDLINE=9920606; PubMed=10192388;			
RA	Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,			
RA	Olinger L., Grimwood J., Davis R.W., Stephens R.S.;			
RT	"Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.";			
RL	Nat. Genet. 21:385-389(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ARJ9;			
RX	MEDLINE=20150255; PubMed=10684935;			
RA	Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,			
RA	White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,			
RA	Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,			
RA	Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,			
RA	Eisen J., Fraser C.M.;			
RT	"Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae A039.";			
RL	Nucleic Acids Res. 28:1397-1406(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=J138;			
RX	MEDLINE=20330349; PubMed=10871362;			
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,			
RA	Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;			
RT	"Comparison of whole genome sequences of Chlamydia pneumoniae J138			

```

RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-183;
RA Geng M.M., Schuhmacher A., Muehlendorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001654; AAD18867.1; -
DR EMBL; AE002165; AAF37914.1; -
DR EMBL; AP002547; BAA98935.1; -
DR EMBL; AE017159; AAP98685.1; -
DR PIR; D72042; D72042.
DR PIR; E86581; E86581.
DR PHCI-2DPAGE; Q927H7; -.
DR TIGR; CP0018; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 651 AA; 68217 MW; 47AE6C3FF2FF0123 CRC64;

Query Match 99.8%; Score 2844; DB 16; Length 651;
Best Local Similarity 99.8%; Pred. No. 2.3e-124;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDYKTAQAYDTIFTSTSLA 60
DB 69 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDYKTAQAYDTIFTSTSLA 128
QY 61 DIQAALVSLQDAVTNIDTAATDEETATAAEWETKNADAVKGAQITELAKYASDNQAIL 120
DB 129 DIQAALVSLQDAVTNIDTAATDEETATAAEWETKNADAVKGAQITELAKYASDNQAIL 188
QY 121 DSLGKLTSPDLLQAALQSVANNKKAELLKEMQDNPVPGKTPAIQAQSLVDQTDATATQ 180
DB 189 DSLGKLTSPDLLQAALQSVANNKKAELLKEMQDNPVPGKTPAIQAQSLVDQTDATATQ 248
QY 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQKFPDPSPI 240
DB 249 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQKFPDPSPI 308
QY 241 LOEAQKVIQAEKDLKNIKPADGSDVPNPGTGGSKQGGSSIGSVMLDDDAENETA 300
DB 309 LOEAQKVIQAEKDLKNIKPADGSDVPNPGTGGSKQGGSSIGSVMLDDDAENETA 368
QY 301 SILMSGFRQMIHMFNTENPDSQAQQLAQAARAAKAGDSDSAALADAKALEALGK 360
DB 369 SILMSGFRQMIHMFNTENPDSQAQQLAQAARAAKAGDSDSAALADAKALEALGK 428
QY 361 AGQQGGILNALGQITASAASVAGVPPAAASIGSVKQLYKTSKSGDYKTOISAGYDA 420
DB 429 AGQQGGILNALGQITASAASVAGVPPAAASIGSVKQLYKTSKSGDYKTOISAGYDA 488
QY 421 YKSINDAYGRARNDATRDVINNVSTPALTRSVPRARTARGPEKTDQALARVIGNSRTL 480
DB 489 YKSINDAYGRARNDATRDVINNVSTPALTRSVPRARTARGPEKTDQALARVIGNSRTL 548
QY 481 GDVYSQVSALQSVQMIIQSNPQANNEIRKQLTSVTKPPQPGFYVQLSNDSTQKFIK 540
DB 549 GDVYSQVSALQSVQMIIQSNPQANNEIRKQLTSVTKPPQPGFYVQLSNDSTQKFIK 608
QY 541 LESLFAEGSRTAAEIKALSFNTSLFIQOVLNVIGLSYGLQ 583
DB 609 LESLFAEGSRTAAEIKALSFNTSLFIQOVLNVIGLSYGLQ 651

RESULT 2
Q46166 PRELIMINARY; PRT; 715 AA.
AC Q46166;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

76 kDa protein.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RA PEREZ-MELGOSA M., KUO C.-C., CAMPBELL L.;
RT "Isolation and characterization of a gene encoding a Chlamydia
RT pneumoniae 76-kilodalton protein containing a species-specific
RT epitope.";
RL Infect. Immun. 62:880-886 (1994).
DR EMBL; L23921; AAA23117.1; -.
DR PIR; I40729; I40729.
SQ SEQUENCE 715 AA; 76626 MW; 3F01C9A59DE964C5 CRC64;

Query Match 65.6%; Score 1870; DB 2; Length 715;
Best Local Similarity 99.7%; Pred. No. 4.4e-79;
Matches 385; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDYKTAQAYDTIFTSTSLA 60
DB 325 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDYKTAQAYDTIFTSTSLA 384
QY 61 DIQAALVSLQDAVTNIDTAATDEETATAAEWETKNADAVKGAQITELAKYASDNQAIL 120
DB 385 DIQAALVSLQDAVTNIDTAATDEETATAAEWETKNADAVKGAQITELAKYASDNQAIL 444
QY 121 DSLGKLTSPDLLQAALQSVANNKKAELLKEMQDNPVPGKTPAIQAQSLVDQTDATATQ 180
DB 445 DSLGKLTSPDLLQAALQSVANNKKAELLKEMQDNPVPGKTPAIQAQSLVDQTDATATQ 504
QY 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQKFPDPSPI 240
DB 505 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQKFPDPSPI 564
QY 241 LOEAQKVIQAEKDLKNIKPADGSDVPNPGTGGSKQGGSSIGSVMLDDDAENETA 300
DB 565 LOEAQKVIQAEKDLKNIKPADGSDVPNPGTGGSKQGGSSIGSVMLDDDAENETA 624
QY 301 SILMSGFRQMIHMFNTENPDSQAQQLAQAARAAKAGDSDSAALADAKALEALGK 360
DB 625 SILMSGFRQMIHMFNTENPDSQAQQLAQAARAAKAGDSDSAALADAKALEALGK 684
QY 361 AGQQGGILNALGQITASAASVAGVPP 386
DB 685 AGQQGGILNALGQITASAASVAGVLP 710

RESULT 3
Q824X3 PRELIMINARY; PRT; 636 AA.
AC Q824X3;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN CCA00015.
OS Chlamydia pneumoniae.
OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RA READ T.D., MYERS G.S.A., BRUNHAM R.C., NELSON W.C., PAULSEN I.T.,
RA HEIDELBERG J., HOLTZAPPEL E., KHOURI H., FEDEROVA M.B., CARTY H.A.,
RA UMAYAM L.A., HAFT D.H., PETERSON J., BEANAN M.J., WHITE O.,
RA SALZBERG S.L., HSIA R.-C., MCCLARTY G., RANK R.G., BAVOIL P.M.,
RA FRASER C.M.;
RT "Genome sequence of Chlamydia pneumoniae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiales.";
```

```
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016994; AAP04768.1; -.
DR TIGR; CCA00015; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 636 AA; 67020 MW; C734944DCE8CA302 CRC64;

Query Match      39.6%; Score 1129; DB 16; Length 636;
Best Local Similarity 42.9%; Pred. No. 9.7e-45;
Matches 252; Conservative 107; Mismatches 212; Indels 16; Gaps 7;

Qy 6 KLGASSSSSSSTSRSDVSTTATPTPPPTDDYKTOAQAYDTTFTSLADICAA 65
Db 7 RLSFLSSARKSLASLFDKISFFSGKT--PQTFDEAKTQAESAKTALQSATTVDDQFKTA 114
Qy 66 LVSLQDAVYNIKDTAATDEETALAEAKETKNADAVKVGQAQITELAKYASDQAIIDSLGK 125
Db 115 LQQLQDAVKQMEQLATTDAEKATVATWKTALAEAKSTLDTNLQNGAILTENQKLEAIKT 174
Qy 126 LTSFLLQALLOSANNKKAELLKEMQDPVVPKTPAIAQSLVDQTDATATQIEKDG 185
Db 175 TSSMDQIMGAAGQVETNKTAAELIKQLKE---AGVSYPVIDDLKQITTSIGTQVTELA 230
Qy 186 NAIRDVAFQGNASGAVENAKSNNSISNDSAKAIAFAKQIAEAQKFPDSPILQAE 245
Db 231 DAISEYAAAGKNSTAAVGAQANNSPANIEASKQTIANAKVIEDALXADPSPILKAAL 290
Qy 246 QMVIQAEKDLKNIKPADGSDVP--NPGT--TVGSKQKQSGSIGSRVSMLLDDAENETAS 301
Db 291 KEQQQAADILNVKPSGSDVPVIGPGAGPSVGTQNRGATLGEVRVSMMLTDDVNETAA 350
Qy 302 ILMGFRMIHMFNTENPDSQAQELAAQAPAAK---AGDSSAAALADAKALEAL 358
Db 351 IIMQFRNMIDFNHDSDFAPLEEIMNQVTDLSQTQINPADAEATAQLQEIQTQIDAL 410
Qy 359 GKAGQGGILNALGOIASAAVVSAGVPPAAASSIGSVKQLYKTSKTS--DYKTKQISA 416
Db 411 QGTAGQGGKINALGALTAAASISTGAPIASANGGSAVKQLYKTSSTAASSKSYADSLA 470
Qy 417 GYDAYKSINDAYGRARNDRVINNVSTPALTRSPVPRARTEARGPEKTDQALARVISGN 476
Db 471 GYGAYQSLNDVYSRS--SASNRVLDRTSTPALTQVTSRTEPRNDNAAQRFARTIAAN 529
Qy 477 SRTLDGVYSQVSAQSVQITQSNFOANNEHROKLTGATVYPPGPPGYVOLSNDSTQK 536
Db 530 SNTLDGVYASVGLQTLGLVQNNFQANEEIKQKLTSEVTKAPQSGYPVHQLSNDSTKK 589
Qy 537 FIAKLESIFAEGSRATAAEIKALSFETNSLFIQQLVNVIGSLYSGYLQ 583
Db 590 FIAQLENEFVQSKRLAEAKAEAFKQPLFIQQLVNVASLFSGYLQ 636

RESULT 4
O84627 PRELIMINARY; PRT; 647 AA.
AC O84627;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CHLPN 76 kDa homolog.
GN CT622.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=97841136;
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans."
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
```

```
DR EMBL; AE001333; AAC68226.1; -.
DR PIR; G71490; G71490.
DR PHCI-2DPAGE; O84627; -.
KW Complete proteome.
SQ SEQUENCE 647 AA; 68525 MW; COD14C2D74473625 CRC64;

Query Match      29.6%; Score 843; DB 16; Length 647;
Best Local Similarity 35.5%; Pred. No. 1.9e-31;
Matches 213; Conservative 113; Mismatches 243; Indels 31; Gaps 14;

Qy 1 MSLADKLGIASSSSSTSRSDVSTTATPTPPPTDDYKTOAQAYDTTFTSL 59
Db 60 ISLRD--AILNKSPTDLSQLEASTSTSVTVAAARDINEAKSNFTAKSGLENATLL 117
Qy 60 ADIQAALVSLQDAVYNIKDTAATDEE--TAIAAEWETKNADAVKYGQAQITELAKYASDQ 118
Db 118 AEYETKMDLMAALQDMERLAKQAEVTRIKEALQEQ---EVDKLNQLVKLEKQNT 173
Qy 119 ILDSIGKLTSPDLLQALLQSVANNKKAELLKEMQDPVVPKTPAIAQSLVDQTDATA 178
Db 174 LKETITTTDSADQIPAINSQLIEINKNSADQIIKLEGQNI---SYEAVLTNAGEVIKASS 230
Qy 179 TQIEKDGNAIRDVAFQGNASGAVENAKSNNSISNDSAKAIAFAKQIAEAOKK--F 235
Db 231 EAGIKLQALQSLVDAGDQSOAVLQAOQNSPNIAATKLLIDAEKYNELKQEHGL 290
Qy 236 PDSPILOEAEQVIOAEKDLKNIKPADGSDVPNPGTTVGSKQKQSGSIGSI-----RV 288
Db 291 TDSLPLVKAERQISOAQKDIQEIKPS--GSDIPIVGPS--GSAASAGSAVGALKSSNNSGRI 348
Qy 289 SMILDDAENETASILMSGFRMIHMFNTENPDSQAQELAAQAPAAK---AGDSSAAA 345
Db 349 SLLEDVDNEMAAATAMQGFREMIQFNVNPNATAKELQAMEAQLTAMSDQLVGDGELPA 408
Qy 346 ALADAQKALEAALGKAGQOQOQILNALGOIASAAVVSAGVPPAAASSIGSVKQLYKT--S 403
Db 409 BIQAIKALQAL--KQSTDLATAMQGVAFAAKAVGGSGAGTAGTVQMVVQLYKTAFS 467
Qy 404 KSTGSDYKTOISAGYDAYKSINDAYGRARNDRVINNVSTPALTRSPVPRARTEARG-P 462
Db 468 STSSSSYAAALSDGVSAYKTLNSLYESRS--GVQSAISOTANPALSRVSRSRGIESQGRS 526
Qy 463 EKTQALARVTSNGSRITLGVYSQVSAQSVQITQSNFOANNEHROKLTGATVYPPGPPQ 522
Db 527 ADASQRAAEITVRDSQTLGVDVYSRLVLDLSMTSTVSNPQVNPQVBEIMQKLTASISKAPQ 586
Qy 523 GYPVYQLSNDSTQKFIKLESIFAEGSRATAAEIKALSFETNSLFIQQLVNVIGSLYSGYL 582
Db 587 GYPAQNSADSLQKFAQLEREFVDGERSLAESRENAPRKQPAFIQQLVNVIASLFSGYL 646

RESULT 5
O8CMU7 PRELIMINARY; PRT; 2310 AA.
AC O8CMU7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Streptococcal hemagglutinin protein.
GN SE2249.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016751; AA005891.1; -.
DR InterPro; IPR008985; Cona-like_lec_gl.
DR InterPro; IPR008009; He_PIG.
DR Pfam; PF05345; He_PIG; 2.
```



```
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003133; BAB42527.1; -.
DR F01; B89921; B89921.
DR InterPro; IPR002988; GA.
DR InterPro; IPR006530; YD.
DR Pfam; PF01468; GA; 46.
DR TIGFAMS; TIGR01643; YD_repeat_2x; 7.
KW Complete proteome.
SQ SEQUENCE 6713 AA; 722339 MW; AF6E8DE226BE4888 CRC64;

Query Match 7.2%; Score 206; DB 16; Length 6713;
Best Local Similarity 21.5%; Pred. No. 1.5;
Matches 145; Conservative 97; Mismatches 279; Indels 154; Gaps 31;

QY 5 DKLGIASNSSSTSRADVDSS---TTATPTPPPTFDYKTAQTAYDTFTS--TSL 59
DB 2215 EKLAARAKQTAQSDIGRLTDLNNAQRTAANAQVDAQPNL-----AAVTAARAKKATSLNTAM 2269
QY 60 ADIQAAVLSLODAVNIK-----DTAAATDEETAIAAEWETKNADAVKYGAQIT 107
DB 2270 GNLKHALAEKONTKRSVNYTDADQPKQAYDVAQAEAITVA--NGSNANETQVQALIN 2327
QY 108 ELAKYASNOAILDSGLKLTSPDLLQALLQSVANNKAAELLKEMQ--DNPVVPQKTPAI 166
DB 2328 QLNQAKND-----LNGDNKVAQAKETAKRALASVNLNNAQSTAATSQIDNATTVADVTA- 2382
QY 167 AGSLVDQDTATQIE---KQGNAI-----RDAYF-AGQNASGAVENAKSN 209
DB 2383 AQNTANELNTAMGQJONGINDQNTVKQVNFDDQGGKQDVAVTNAQGLDXANGON 2442
QY 210 SISNIDSAAALATAKTAQIAERAKKFPDPSPIQEAQVIAQEKD-----LKNIK 259
DB 2443 -----MYKAQVEALNQVTTAKNAL-----NGDANVRQAKSDAKANGLTLHLNNAQ 2489
QY 260 PADGSDVNPPTGVG-----SKQGG-----SSIGS-----IRVSMLLDQAEKETASI 302
DB 2490 KQDLTSQIEGATTVNGVSVKTKAQDLGAMQRLSEAIANKQDTRASENYIDADPTKKA 2549
QY 303 LMSGFROMHMENTE---NPDQQAQOELAAQARAQAA--GDSAAALADRAQALEAL 358
DB 2550 FDNALTAESYLNKHGNTKQKQAEQVQSTSTENALNGDANIQCAKTEATQALDNL 2609
QY 359 GKAGQOQGLNALGOTASAAVVSAGVPPA--AASIGTSVKQLYKTSKSTGSDYKTOISA 416
DB 2610 QLNTPOK---TALKQOVNAAQVSGVTDLKNATSILNNAMDQJ---KQAGIDHDTIVAG 2662
QY 417 GY-----DAYKSINDAYGRAN--DATRDVINNVIS--TPALTRSPVPARTTEARG----- 461
DB 2663 GNYTNASPDKQAYTDAYNAAKNIYNGSPNVTNAADVTAATQRVNNAETSILNGDTNLAT 2722
QY 462 -PEKTDQALARY-----ISGNSRLTGDVYVSQVSALQSMQITOS-----NP 501
DB 2723 AKQKADLARQWTHLSDAKQKQITGOI-----DSATCVTGQVSKVONANLNDANNQLRNS 2778
QY 502 QANNEEIRQKLTSAVTKPPFGFYVOLNSDSTQKF---IAKLSLFAEGSRTAAEIKAL 558
DB 2779 IANKDEVKA-----SQPYVDADTDQKNAYNTAVTSAENIINATSOPTLDPSAV 2826
QY 559 SFETNSLFTQQVLVN 573
DB 2827 TQAAQNVNTKTNALN 2841

RESULT 8
Q07290
ID Q07290 PRELIMINARY; PRT; 1822 AA.
AC Q07290;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
```

```
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Epf* protein.
GN EPF*.
OS Streptococcus suis.
OC Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1307;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1890;
RX MEDLINE=93328288; PubMed=9335363;
RA Smith H.E.; Reek F.H.; Vecht U.; Gielkens A.L.J.; Smits M.A.;
RT "Repeats in an extracellular protein of weakly pathogenic strains of
RT Streptococcus suis type 2 are absent in pathogenic strains.";
RL Infect. Immun. 61:3318-3326(1993).
DR EMBL; X71880; CAA50714.1; -.
DR PIR; S33441; S33441.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_YSIIRK.
DR Pfam; PF04650; YSIIRK signal; 1.
DR TIGFAMS; TIGR01168; YSIIRK signal; 1.
SQ SEQUENCE 1822 AA; 192632 MW; 3838960C77641D7D CRC64;

Query Match 7.2%; Score 205; DB 2; Length 1822;
Best Local Similarity 23.7%; Pred. No. 0.3;
Matches 145; Conservative 82; Mismatches 229; Indels 156; Gaps 27;

QY 37 PTFDDYKTKQAQ-----TAYDTFTSTSLADIQAALVSLQDA--VTNIKDTAA 81
DB 1265 PNLSDQEQAEIKKLTDAVAKTLATIRDNADKRTQEAKEKAQALADLEKAKETQKIADKAA 1324
QY 82 TDEETAI--AEWETKNADAVKVGQAITELAKYASDNOAILDSGLKLTSPDLLQALLQ 139
DB 1325 IDRLTILVKGELEATKQDAKNKIATAKAAAAKEAATASNPNLTDAEKKTFTDAVDA---E 1380
QY 140 VANNKA-----AELLKENQDNVVPVPGKTPATAQSLVDQTDATA--TQIEKGNAIRDA 191
DB 1381 VAKANDAI:SAATSPADVQKE--EDAGVA-----ATAEDVLDAAKQDAKNKIATAKAAAAKEA 1434
QY 192 YFAGNAGSA-----VENAKSNNSISNIDSAKAAIATAKTQIAEAQKFPDPSPI 241
DB 1435 IGSNPNLTDAEKKTFTDAVDAEAVAKANDAIIS-----AATSPADVQKE--EDAGVA 1482
QY 242 QEAQMTVQAEKDKLNKIKPADGSDVNPGETTVGSGKQGGSSIGSIRVSMLLDDAENETAS 301
DB 1483 AIAEDVLDAAKQDAKN--KIAKESD-----AAKSAIDANPNLTDAEKESAK 1526
QY 302 ILMGSRQMIHMFENTENPDSQAQOQELAA-----QARAAKAGDSDSAAALADAKQ--- 352
DB 1527 -----KAVDADAKAATDAIDASTSPVEQAEDKGVGSIQDVLDAAKQDAK 1573
QY 353 ---ALEALGKA-----GQQQGLNALGOTASAA--VVSAGVPPAAASS-----I 392
DB 1574 NKIAKEVAARAEADANPNLSDAEAKSKAVDADAKATTDAIDASTSPVEQAEDKGV 1633
QY 393 GSSVKQLYKTSKSTGSDYKTOISAGYDAYKSIDNAYGR---ARNDATRDVIN---NVSTP 446
DB 1634 GSIRQDVLDAAK--QDAKNKIAKESDAKSAIDANPNLTDAEKESAKKAVDADAKAATD 1690
QY 447 ALTRSPVPARTTEARGPE--KTDQALARVI-----SGNSRTLGDVYVSQVSALQSMQITOS 499
DB 1691 AIDAST--SPVEQAEDKGVGATKDLIDAAKQDAKNKIATAKESAKSVISDSPNLTDA 1748
QY 500 NPQANNEEIRQKLTSAVT-----KPPQFGFYVOLNSDSTQKF 537
DB 1749 AKEAKSEIDKAVEAIVLINGVETQYELEKIKLPMAALIKPAKAVTPVVD--PNNLTEKE 1807
QY 538 IAKLESFPAEGS 549
DB 1808 IARIKATFLKENN 1819

RESULT 9
```



```
Db 2663 GNYTNASDKGAYTDAYNAAKNIIVNGSPNITNAADVTAATQVRVNNNAETSLNGDTNLAT 2722
Qy 462 -PEKTDQALARV-----ISGNSRTLGDVYSQVSALQSVNQITQS-----NP 501
Db 2723 AKQQAQALRWTHLSDAQKQSIQTQI-----DSATQVTGVQSVKDNATNLNANQNLNS 2778
Qy 502 QANNEIRKQKITSATVPKPPQPGYPYVQLSNDSTOKF-----IAKLSLFAEGSRTAAEIKAL 558
Db 2779 IANKDEVKA-----SOPYVDADTDKQAYNTAVTSAENIINATSCPTLDRSAV 2826
Qy 559 SFETNSLFIQVIVN 573
Db 2827 TQANQVNTKALN 2841

RESULT 11
Q8CP76 PRELIMINARY; PRT; 9439 AA.
AC Q8CP76;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ebha protein.
GN SE1128
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016747; AA004725.1; -.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR002988; GA_B.
DR Pfam; PF05738; Cna_B; 1.
DR Pfam; PF01468; GA; 39.
KW Complete proteome.
SQ SEQUENCE 9439 AA; 1050771 MW; 5A8C68E9AA31CDD4 CRC64;

Query Match
Best Local Similarity 7.0%; Score 200.5; DB 16; Length 9439;
Matches 132; Conservative 99; Mismatches 255; Indels 161; Gaps 27;

Db 12 SNSSSTSRSDVSTTATATPTPPPTDDY-----KTQATAYDTIFTSTSLAD 61
1945 NNGDATSQINDAKNTVEQAKR-----DYVEAKSNLRADKSQLQASDYTLNRDVLTD 1997

Qy 62 IQAALV-SLQDAVNIKDTATDEETAAAEWETKNADAVKVGQITELAKYASDNQAIL 120
1998 KKPASVRRYNEAISNIR-----KELDTAKADASST-LRNTNPSVEQVR 2039

Qy 121 DSLGKLTSTF--LLQA--ALLOSVAANNKAAELKEMQDN-----PVVPGKTP----- 164
2040 DALANKINTVQKVAQIALLPKNNSELVQAKKELQDAVNDIPOTQGTQOTINNYNDK 2099

Qy 165 -----ALAQSLVDQTDATATQIEKDGNAIRDAYFAGQNASG-----A 201
2100 QREABRALTSARQVINDGDAITQIEITSEKSKVEQAMQALTNAKSNLRADKNELOQAYNKL 2159

Qy 202 VENAKSN--SISNIDSAKAAIATAKTOIAEAKKFFDPSPILOAEQMWIAEKDLKN 257
2160 IENVSTNGKKPASTROYETAKARI---QNQINDAKNE--AERILGNDNPQVSQVTOALNK 2214

Qy 258 IKPADGSDVPNPGTIVGSGSQGSSIGSIRVSMILDDDAENETASILMSGFQWIMHMENTE 317
2215 IKAIQ---PKLTEAINLMQKNENTELINAKRLNVAVNDTDP--THGMTQ--ETINNY 2266

Qy 318 NPDSQAQOEIAAQARAAKAGDDSAALADAQKALEAALGKAGQOQGIILNALGQIASA 377
2267 NAKKREAQNEI--QKANMIINNGDATAQDISSEKSKVEQVL-----QALQNAKNDLR-- 2316
```

```
Qy 378 AVWSAGVPPAAAGSIGSSVKQLYKTSKSTG-----SDYKTIQISAGDYAYKSINDAYGRA 431
Db 2317 -----ADKRELQATAYNKLQNVNTNGKPPSIQNYKS-----ARRNIENQYNTA 2360
Qy 432 RNDA-----TRDVIN-----NVSTPALTFESVPRARTE-----ARGPEKTDQALA 470
Db 2361 KNEAHNVLENTPTVNAVEDALRKINAIQPEVTKAINILOKEDKNSSELVRAKELKDRAIN 2420
Qy 471 RVISGNSRITLGDVYSQVSALQSVNQITQS-----NPOANNEETR-----QKLTSAVTK 518
Db 2421 SQPSLNCMTQESINNYTTKREAQNIASSADTIINNGDASIEQITENKIRVEEATNALNE 2480
Qy 519 PPQFGYPYVQLSNDSTOKFIAKJESLFAEGSRTAAEIKALSFETNSL 565
Db 2481 AKQ-----HLTADTTSLKTEVRKLRRGDTNNKKPSSVSAYNNTIHSL 2523

RESULT 12
Q9NE28 PRELIMINARY; PRT; 1313 AA.
AC Q9NE28;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN L232.01.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Klages S., Borzym K., Beck A., Reinhardt R., Ivens A.C., Quail M.,
Rajandream M.A., Barrell B.G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RX Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; ALJ54553; CAB89628.2; -.
KW Hypothetical protein.
SQ SEQUENCE 1313 AA; 140188 MW; 2BA3034A75AD23AC CRC64;

Query Match
Best Local Similarity 7.0%; Score 200; DB 5; Length 1313;
Matches 126; Conservative 97; Mismatches 251; Indels 88; Gaps 25;

Qy 28 TATATPTPPPTDDYKTAQATAYDTIFTSTSLADIQAA-LVSLQDAVNTIKDTAATDEET 86
244 SSTTPTPPSS-----DMAYQLTRLEBSTAHPEDLQRRITAVVASRDAPAVT 291

Qy 87 AIAAEWETKNADAVKVGQITELAKYASDNQAILDSLGLKLTSPDLLQAALLOSVAANNKA 146
292 SAPAELSTAVEAVQ-----KAAAEQ-----EKLNLINDLLAEERLAARGSGGDA 339
147 AELLKEKQDNPNVPGKTPAJAQSLVDQTDATAT-----QIEKDGNAIRDAYFAGQ--NASG 200
340 ANAIVDMR-----ATBESTESAVQSAINELRRTITSEVKKLEK-SKGTGEVTAGSATDADM 395
201 AVENAKSNNSISNIDSAKAAIATAKTOIAEAKKFFDPS-PILOAEQMWIAEKD-LKNI 258
396 AIEIEKVFDAVQSQSLAAVDEVVGGIFKEQAATRDALRVQDE----VLQRLRDELGRGT 451
259 KPADG-SDVNPNGTT-VGSGKQSGSSIGSTRVSMILDDDAENETASILMSGFQWIMHMENT 316
452 VATEATASTPSPAAPVNGVPEAAQQ---KITEVLQRLRRD-----VSATSSMLEIHTND 502
317 ENPDSQAQOEIAAQARAAKAGDDSAALADAQKALEAALGKAGQOQGIILNALGQIAS 376
```

```

Db 503 N-----AQOQMT---VAQWSDSVTAALQOHTTEALQSAMTTVOBQLNALREAGEAAS 553
QY 377 AAVVSAGVPPAAASSGSSVKQLYKTSKSGDYKTOISAGYDAYKSINDAYGRANDAT 436
Db 554 AASAAPPAPAPSPAPLSEVLTDVAEVLSPRWAEISV---QVEKNNTANKLAIEQQL 610
QY 437 ROVINNVST---PALTRSVPRARTEARGPEKTPQALARVIGNSRSLGDIYSQVS---ALQ 491
Db 611 LHIADSVSTVAASLOAQLEEMRTSABSQLKTEHA-----AKEQGSVIDTSPAATA 662
QY 492 SVMQITQSNPOANNEIROKLTSAVTKPPQFG-----YPYVQLSNDSTQKFIKLESIFAE 547
Db 663 ALIEKTHEATTREELQALQAOQVGKTPENVLTPMYKYI---DSVLTFMK--EELSLO 716
QY 548 GSRTAAEIKALSPETNSLFQIQ 569
Db 717 ETNLTKAMASLA-ETVTAAVRE 737

RESULT 13
QSNWQ6
ID QSNWQ6 PRELIMINARY; PRT; 9904 AA.
AC QSNWQ6
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Ebn protein.
GN EBN OR MW1324.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004826; BAB95189.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002988; GA.
DR InterPro; IPR005877; Gpos_YSRK.
DR InterPro; IPR006530; YD.
DR Pfam; PF01468; GA; 48.
DR Pfam; PF04650; YSRK signal; 1.
DR TIGRFAMs; TIGR01643; YD repeat 2x; 8.
DR TIGRFAMs; TIGR01168; YSRK_signal; 1.
KW Complete proteome.
SQ SEQUENCE 9904 AA; 1068476 MW; CB86908F75EA5682 CRC64;

Query Match 7.0%; Score 200; DB 16; Length 9904;
Best Local Similarity 20.8%; Pred. No. 4.6;
Matches 143; Conservative 103; Mismatches 262; Indels 180; Gaps 29;

QY 5 DKIGIASSNSSSTSRGADVSTTATPTPPPTDDYKTOAQPAYDTFTSTSLADIOA 64
Db 5483 EKLAARQAKTSKDTGRULDLNNAQRTA-----ANAEVDQAPLAAVTA 5525
QY 65 A---LVLSIQAVTNKTAATDEETAIAAEWETKNADAVKVGQITELAKYASDNQAILD 121
Db 5526 ANKATSLNTANGNLKHALAEKONTKGSVNY--TDADQPKQOAVDTAV---TQAEAITN 5579
QY 122 SLGKLTDFDLQALLO-----SVANNKKAELLKEM 153
Db 5580 ANGSAENETTVOALNQLNQAENDLNGDNKVAQAKESAKALASYNLNAQAQSTAAATSOI 5639
QY 154 QDNVPVPGKTPAIQSLVDQDTATQIE---KGNAI-----RDAYF-AGQ 196
Db 5640 DNATTVAGVT--AAQNTANELNTAMGQLQNGINDQNTVKQOVNFTDADQGGKDAYTNVAT 5697
QY 197 NASGAVENAKSNNSINSDSKAALATAKTQIAEAKQKFPDPSPILOEAEQWVQAKD-- 254

```

```

Db 5698 NAQGLDKANGQN-----MTKAQVEAALNQVTTAKNAL-----NGDANVRQAKSDAK 5744
QY 255 -----LKNIKPADGSDVPNPPTVGG-----SKQOG-----SSIGS---IRVS 289
Db 5745 ANLGTLTTHLNNAKQDLTSCIEGATTVNGVSVKTKAQLDGMORLESIAINKOQTKAS 5804
QY 290 MLDDDAENETASTILMSGFGRMIHMENTE---NPDSQAQOQELAAQAAAKAA--GDSAAA 345
Db 5805 ENYIDADPTKKTAFDNTAITQAESYLNKDHGTNKQAVQQAICSVTSTENALNGDANLQ 5864
QY 346 ALADAQKALEALGKAGQOQGITNALGQIASAAVSVAGVPPA--AASSTGSSVKQLYKTS 403
Db 5865 AKTEATQATDNLTHLNTPKQ--TALKQOVNAAQRVSGVTDLKNASATSLDNAMDQ--- 5917
QY 404 KSTGSDYKTIISAGY-----DAYKSINDAYGRARN--DATRDVINNVV--TPALTESVPR 454
Db 5918 KQGIADHDITVAGGNVTNASPKQGGAYTDAYNAKNIVNGSENVITNAADVTAARTORVNN 5977
QY 455 ARTEARG-----PEKTDQALARV-----ISGNSRTLGDVYSQVSALQSVMOIT 497
Db 5978 AETSLNGDSNLATAKQAKDALRQHTLSDAQKQSITGOI-----DSATQVTGVQSVKDNA 6033
QY 498 QS-----NPOANNEIROKLTSAVTKPPQFGYPYVQLSNDSTQKF---IAKLESIF 545
Db 6034 TNLDNAMNQLRNSIAKDEVKA-----SQPYVDADTDKQNTAYNTAVTSAENII 6081
QY 546 AEGSRTAAEIKALSPETNSLFQIQVLN 573
Db 6082 NATSQTLDPISAVTQANQVNTKNTALN 6109

RESULT 14
QSNWU3
ID QSNWU3 PRELIMINARY; PRT; 2275 AA.
AC QSNWU3
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE MW2575 protein.
GN MW2575.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004831; BAB96440.1; -.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR008009; He_PIG.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF05345; He_PIG; 2.
DR PROSITE; P050847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 2275 AA; 228407 MW; E676B6BB9E60CE2E CRC64;

Query Match 7.0%; Score 199.5; DB 16; Length 2275;
Best Local Similarity 19.3%; Pred. No. 0.73;
Matches 115; Conservative 127; Mismatches 235; Indels 119; Gaps 22;

QY 2 SLAD----KLGIASSNSSSTSRGADVSTTATPTPPPTDDYKTOAQATVDTFTSTIS 58
Db 1236 SLSDSTSESGSTSTLSLNSSTGSASISTSTISEST-----STFKSESVSTSLSMSTIS 1290
QY 59 LADIQAALVSLQAVTNIK-DTAATDEETAIAAEWETKNADAVKVGQAIT-ELAKYASDN 116

```

```
Db 1291 LSDSTLSLSDSTSDSKSDSLSTSMSTSDSI--STSKSDSISLSTSLSGSTSESESDS 1348
Qy 117 QAILDSLGLKSLFDFLLQALLOQSVANNKKAELLKEMQNPVPGKTPAIAQSLVDQIDA 176
Db 1349 TSSSEKSDSTSMSTI--SMSTSGSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 1384
Qy 177 TATQTEKDNRAIDAYFAGQONAGAVENAKSNNSISNIDSAKAAIATAKTOIAEAKKFP 236
Db 1385 TSTSL-----SLSAMNQSG-VDSNSAQSSASN-----STSTSTSESDSQST 1425
Qy 237 DPTIIOEAEQVIOAEKDLKNIPKADGSDVPNPGTIVGSGKQSGSIGSIQSVMLLDDAE 296
Db 1426 SG-----YTSQSTSQSESTSTSTSLSD-----STSIKSTSQSGSV-STASLSGSESE 1473
Qy 297 NETASLISGFRQMHMFNTENPDQOAAQCELAQ-----ARAAGAAGDSDAAAAALADAQ 351
Db 1474 SDPSQISLTSASSTSESASLSDSTSTSNSSGASTSTSLNSASASSESDSSSTSLSDST 1533
Qy 352 KA-LEALGKACQOQIGL-----NALGOIAAAVVSAGVPPAAASSIGSSVKQLYK 401
Db 1534 SASMQSSSDSQSTGASLSDSTSTSNRMSTIAS-----LSTSVSTSESGSTSESTSESDS 1590
Qy 402 TSKSTGSDYKT--QISAGVDAYKSIDAYGRANDAT---RDVINNVSTPALTSVPRA 455
Db 1591 TSTSLSDSQSTGRSTASAGSASTSTSTSDRSTASSTSTSMRTSTSDSQSMSTSTSTSTST 1650
Qy 456 RTEARG-----PEKTDQALARVISGN--SRTLGVDVYSQVSLQSVMQITQNPQANNEE 507
Db 1651 MSDSTSLSDSVSDSTSDSTASSTSGMSVSIISLSDSTSTSTSEVMSASISDSQSMSES 1710
Qy 508 IRQKLTSAVTKPPQPGYPIVQLSNSTQKFIKLSLFAEGRTAEIKALSFETN 563
Db 1711 VN-----DSESVSESNSESDSKSMGSGTSTVSDSGSLSVST 1746

RESULT 15
Q8XDQ4
ID Q8XDQ4 PRELIMINARY; PRT; 973 AA.
AC Q8XDQ4;
DC Q8XDQ4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative membrane protein of prophage CP-933X (Putative tail fiber protein).
DE GN Z1918 OR ECS1650.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
RN [3]
RP EMBL; AB005333; AAG56007.1; --
```

```
DR EMBL; AP002555; BABJ5073.1; ALT_INIT.
DR PIR; B90835; B90835.
DR PIR; C85693; C85693.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR008969; CarboxypepD_reg.
DR InterPro; IPR005003; Phage_fiber.
DR InterPro; IPR005068; Phage_fiber_2.
DR Pfam; PF03335; Phage_fiber; 6.
DR Pfam; PF03406; Phage_fiber_2; 3.
KW Complete proteome.
SQ SEQUENCE 973 AA; 96316 MW; 776580D2A87E1B36 CRC64;

Query Match 7.0%; Score 199; DB 16; Length 973;
Best Local Similarity 20.6%; Pred. No. 0.26;
Matches 133; Conservative 79; Mismatches 227; Indels 206; Gaps 20;

Qy 9 IASNSSSSTSRSDVD-----STTATAPTPTPP-----TTDD 41
Db 40 VASENPDEAGRYSDVEYQYSVILLVEGFPSPHAGTIVYEDSQGLNDFLGWTEDD 99
Qy 42 YKTAQTAAYDTIF-----TSTSLADIQAALVLSQDAVTNIKDTAATDEETAIAAEWTK 95
Db 100 VEPEALRRFELWVEEVARNSAVAQNTAAAKKSADASTSARE-----AATHATD 149
Qy 96 NADAVKGAQITELAKYASDNOAILDSLGLTSPDLLQALLOQSVANNKKAELLKEWQD 155
Db 150 AADSAR--AASTSAGQAASSAQSASSAGTAST----- 180
Qy 156 NPVVGKTPAIAQSLVDQTATATQIEKDGNAIRDAYFAGQONAGAVENAKSNNSISNID 215
Db 181 -----KATEASKSAASAAESSKSAATSAAGTSETNAVSQ-Q 218
Qy 216 SAKAAIATAKTQIAPAQKFPDPILQBAEQVIOAEKDLKNIKPADGSDVPNPOTTTVG 275
Db 219 SAATSASTATTKASEAASSARDASAKEA-----AKSSETSAASSASSAASS 265
Qy 276 SKQGSISGIRSVSMLLDDAENETASILMSGFRQMHMNTENPDQOAAQCELAQARAA 335
Db 266 ATAAGNSAKAAKTS-----ETNAKSETAAEQSAS 295
Qy 336 KAAGDSDSAAALADAQAALGKAGQOQGLNALGOIAAAVVSAGV-----P 385
Db 296 AAGSKTAALSA-----SAASTSAGQASASATAAGKSAESAASSASTATTKAGEATEQ 349
Qy 386 PAAASSIGSSVKQL---YKTSKSTGSDYKTOISAGYDAYKSINDAYGRANDATR--DVI 440
Db 350 ASAAASSAAKTSSTNAKASSETSAESSKTAASASSASSASSASSASASASAKDEATQASAA 409
Qy 441 NNVTSTPALTSVPRA-----RTEARGPEKTDQALARVISGNSRTLGD 482
Db 410 KSSATTASTKATEAAGSATAAAQSKSTAESAATRAETAAKRAEDIASAVALDEDATTKKG 469
Qy 483 VYSQVSALQSVNQITQSNPQ-----ANNEEIRQKLTSVATKPPQFGYPVVLSEN--D 532
Db 470 IVQLSSAINTSSESLAATPKAVKAAVELANGKYTAQDATT-----QXG--IVQLSNATN 522
Qy 533 STQKFI-----KLESIFAGSRTFAE-----IKALSFTNS 564
Db 523 STSEMLAATPKSVKAAAYDLANGKYTAQDATTQAQKIVQLSSATNS 567

Search completed: March 24, 2004, 05:59:12
Job time : 46.9152 secs
```